

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

CM nucleic - nucleic search, using sw model

Run on: November 15, 2003, 11:42:01 ; Search time 3702.44 Seconds
(without alignments)
20441.655 Million cell updates/sec

Title: US-09-671-050-11
Perfect score: 945
Sequence: 1 atggaagaagtatgaaaaatt.....aggtacttcgcgtcaaaaagt 945

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : GenEmbl.*
- 1: gb_ba.*
 - 2: gb_htg.*
 - 3: gb_in.*
 - 4: gb_om.*
 - 5: gb_ov.*
 - 6: gb_pat.*
 - 7: gb_ph.*
 - 8: gb_pl.*
 - 9: gb_pr.*
 - 10: gb_ro.*
 - 11: gb_sts.*
 - 12: gb_sy.*
 - 13: gb_un.*
 - 14: gb_vi.*
 - 15: em_ba.*
 - 16: em_fun.*
 - 17: em_hum.*
 - 18: em_in.*
 - 19: em_mu.*
 - 20: em_om.*
 - 21: em_or.*
 - 22: em_ov.*
 - 23: em_pat.*
 - 24: em_ph.*
 - 25: em_pl.*
 - 26: em_ro.*
 - 27: em_sts.*
 - 28: em_un.*
 - 29: em_vi.*
 - 30: em_htg_hum.*
 - 31: em_htg_inv.*
 - 32: em_htg_other.*
 - 33: em_htg_mus.*
 - 34: em_htg_pln.*
 - 35: em_htg_rod.*
 - 36: em_htg_mam.*
 - 37: em_htg_vrt.*
 - 38: em_sy.*
 - 39: em_htgo_hum.*
 - 40: em_htgo_mus.*
 - 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	945	100.0	945	6	AX107722	AX107722 Sequence
2	929	98.3	972	6	AX107716	AX107716 Sequence
3	868.8	91.9	1678	6	AX698820	AX698820 Sequence
4	863	91.3	1083	6	AX166534	AX166534 Sequence
5	839	88.8	1041	6	AX107720	AX107720 Sequence
6	823	87.1	1068	6	AX107714	AX107714 Sequence
7	647	68.5	687	6	AX286069	AX286069 Sequence
8	647	68.5	882	6	AX286067	AX286067 Sequence
9	643	68.0	911	6	AX056404	AX056404 Sequence
10	559.2	59.2	2615	6	AX056405	AX056405 Sequence
11	507	53.7	1819	6	AX107724	AX107724 Sequence
12	454.4	48.1	561	6	AX107712	AX107712 Sequence
13	454.4	48.1	594	6	AX107718	AX107718 Sequence
14	415.4	44.0	1612	6	AX399975	AX399975 Sequence
15	413.8	43.8	1363	9	HSSTHPKB	X66358 H.sapiens m
16	400.2	42.3	2952	3	AK112442	AK112442 Ciona int
17	287	30.4	1701	6	AR221995	AR221995 Sequence
18	287	30.4	3080	6	AR221994	AR221994 Sequence
19	287	30.4	3153	4	AB029045	AB029045 Oryctolag
20	283.8	30.0	1993	9	HSU35146	U35146 Homo sapien
21	272.6	28.8	1518	10	AB029067	AB029067 Mus muscu
22	272.6	28.8	3988	10	AB029066	AB029066 Mus muscu
23	272.6	28.8	4022	10	AB029065	AB029065 Mus muscu
24	227.4	24.1	1513	9	AF130372	AF130372 Homo sapi
25	227.4	24.1	1773	6	AX166533	AX166533 Sequence
26	227.4	24.1	2146	9	BC041799	BC041799 Homo sapi
27	222.6	23.6	1620	10	BC028871	BC028871 Mus muscu
28	218.4	23.1	101340	2	AC018104	AC018104 Drosophil
29	218.4	23.1	190668	3	AC007419	AC007419 Drosophil
30	218.4	23.1	260731	3	AE003611	AE003611 Drosophil
31	211.4	22.4	1732	10	AF112183	AF112183 Rattus no
32	211.4	22.4	1738	10	AF112184	AF112184 Rattus no
33	208.4	22.1	1974	3	AF326965	AF326965 Trypanoso
34	200.4	21.2	80362	9	AC079615	AC079615 Homo sapi
35	194	20.5	3399	9	AY217744	AY217744 Homo sapi
36	194	20.5	3399	9	HSTXP3	Y15057 Homo sapien
37	194	20.5	3441	9	BC036091	BC036091 Homo sapi
38	180.2	19.1	3321	3	LME293284	AJ293284 Leishmani
39	175.8	18.6	1866	6	AX056407	AX056407 Sequence
40	171.8	18.2	903	8	AF026545	AF026545 Pneumocys
41	171.4	18.1	2982	9	AF390028	AF390028 Homo sapi
42	168	17.8	960	3	TCU69958	U69958 Trypanosoma
43	163.6	17.3	921	3	DDICRP	L00652 Dictyosteli
44	163.4	17.3	1476	6	AR087351	AR087351 Sequence
45	163.4	17.3	1476	6	AR281916	AR281916 Sequence

ALIGNMENTS

RESULT 1

AX107722

LOCUS AX107722 945 bp linear PAT 30-APR-2001

DEFINITION Sequence 11 from Patent WO0123579.

ACCESSION AX107722

VERSION AX107722.1 GI:13923202

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and Sands,A.T.

TITLE Human kinase proteins and polynucleotides encoding the same

JOURNAL Patent: WO 0123579-A 11 05-APR-2001;
Lexicon Genetics Incorporated (US)
Location/Qualifiers
FEATURES source
1. .945
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 323 a 165 c 202 g 255 t
ORIGIN
Query Match 100.0%; Score 945; DB 6; Length 945;
Best Local Similarity 100.0%; Pred. No. 8.6e-206;
Matches 945; Conservative C; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGAAAAGTATGAAAAATTAGCTAAGACTGGAGAAGGGTCTTATCGGGTTGTATTCAA 60
Db 1 ATGGAAAAGTATGAAAAATTAGCTAAGACTGGAGAAGGGTCTTATCGGGTTGTATTCAA 60
Qy 61 TGCAGAAACAAACCTCTGGACAAGTAGTAGCTGTTTAAAAAATTTGTGGAATCTGAAGAT 120
Db 61 TGCAGAAACAAACCTCTGGACAAGTAGTAGCTGTTTAAAAAATTTGTGGAATCTGAAGAT 120
Qy 121 GATCCTGTTTAAAGAAAATAGCACTAAGAGAAAATACGTAATGTTGAAGCAATTTAAACAT 180
Db 121 GATCCTGTTTAAAGAAAATAGCACTAAGAGAAAATACGTAATGTTGAAGCAATTTAAACAT 180
Qy 181 CCAAAATCTTGTGAACCTCATCGAGGTGTTTCAGAGAAAATGCAATTTAGTTT 240
Db 181 CCAAAATCTTGTGAACCTCATCGAGGTGTTTCAGAGAAAATGCAATTTAGTTT 240
Qy 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT 300
Db 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT 300
Qy 301 GGAGTGATCAAAACCGTATTATGGCAAAACACTTCAAGCTCTTAAATTTCTGCATATACAT 360
Db 301 GGAGTGATCAAAACCGTATTATGGCAAAACACTTCAAGCTCTTAAATTTCTGCATATACAT 360
Qy 361 AACTGTATTCACAGAGATATAAAACCTGAAATAATTTCTAATAACTAAGCAAGGAATAATC 420
Db 361 AACTGTATTCACAGAGATATAAAACCTGAAATAATTTCTAATAACTAAGCAAGGAATAATC 420
Qy 421 AAGATTTGTGACTTCGGGTTTGCACAAATTTCTGATCCAGGAGATGCCATACACCGATTAT 480
Db 421 AAGATTTGTGACTTCGGGTTTGCACAAATTTCTGATCCAGGAGATGCCATACACCGATTAT 480
Qy 481 GTAGCTACGAGATGGTACCGAGCTCCTGAACCTTCTTGTGGGAGATACCTCAGTATGGTTCT 540
Db 481 GTAGCTACGAGATGGTACCGAGCTCCTGAACCTTCTTGTGGGAGATACCTCAGTATGGTTCT 540
Qy 541 TCAGTCGATATATCGGCTATTGGTTGTTTTCAGAGCTCCTGACAGGCCAGCCACTG 600
Db 541 TCAGTCGATATATCGGCTATTGGTTGTTTTCAGAGCTCCTGACAGGCCAGCCACTG 600
Qy 601 TGGCCTGGAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGAAAATTA 660
Db 601 TGGCCTGGAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGAAAATTA 660
Qy 661 ATCCCAAGACATCAATCAATCTTTAAAGTAACGGGTTTTTCCATGGCATCAGTATACCT 720
Db 661 ATCCCAAGACATCAATCAATCTTTAAAGTAACGGGTTTTTCCATGGCATCAGTATACCT 720
Qy 721 GAGCCAGAAGACATGGAAACTCTTGAGGAAAAGTTCTCAGATGTTTCCTGTGGCTCTG 780
Db 721 GAGCCAGAAGACATGGAAACTCTTGAGGAAAAGTTCTCAGATGTTTCCTGTGGCTCTG 780
Qy 781 AACTTCATGAAGGGTGTTCTGAAGATGAATCCAGATGACAGATTAACTGTTCCCAACTC 840
Db 781 AACTTCATGAAGGGTGTTCTGAAGATGAATCCAGATGACAGATTAACTGTTCCCAACTC 840
Qy 841 CTGGAGAGCTCCTACTTTGATTCTTTTCAAGAGGCCCAAAATTTAAAGAAAAGCACGTAAT 900
Db 841 CTGGAGAGCTCCTACTTTGATTCTTTTCAAGAGGCCCAAAATTTAAAGAAAAGCACGTAAT 900

Qy 901 GAAGGAAGAAACAGAAAGACGCCAACAGGTACTTCCGCTCAAAAAGT 945
Db 901 GAAGGAAGAAACAGAAAGACGCCAACAGGTACTTCCGCTCAAAAAGT 945
RESULT 2
AX107716
LOCUS AX107716 972 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 5 from Patent WO0123579.
ACCESSION AX107716
VERSION AX107716.1 GI:13923199
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and
Sands,A.T.
TITLE Human kinase proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0123579-A 5 05-APR-2001;
Lexicon Genetics Incorporated (US)
FEATURES Location/Qualifiers
source 1. .972
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 323 a 172 c 214 g 263 t
ORIGIN
Query Match 98.3%; Score 929; DB 6; Length 972;
Best Local Similarity 100.0%; Pred. No. 3.9e-202;
Matches 929; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGAAAAGTATGAAAAATTAGCTAAGACTGGAGAAGGGTCTTATGGGGTTGTATTCAA 60
Db 1 ATGGAAAAGTATGAAAAATTAGCTAAGACTGGAGAAGGGTCTTATGGGGTTGTATTCAA 60
Qy 61 TGCAGAAACAAACCTCTGGACAAGTAGTAGCTGTTAAAAAATTTGTGGAATCTGAAGAT 120
Db 61 TGCAGAAACAAACCTCTGGACAAGTAGTAGCTGTTAAAAAATTTGTGGAATCTGAAGAT 120
Qy 121 GATCCTGTTTAAAGAAAATAGCACTAAGAGAAAATACGTATGTTGAAGCAATTTAAACAT 180
Db 121 GATCCTGTTTAAAGAAAATAGCACTAAGAGAAAATACGTATGTTGAAGCAATTTAAACAT 180
Qy 181 CCAAAATCTTGTGAACCTCATCGAGGTGTTACAGAGAAAAGGAAAATGCAATTTAGTTT 240
Db 181 CCAAAATCTTGTGAACCTCATCGAGGTGTTACAGAGAAAAGGAAAATGCAATTTAGTTT 240
Qy 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT 300
Db 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT 300
Qy 301 GGAGTGATCAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAAATTTCTGCATATACAT 360
Db 301 GGAGTGATCAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAAATTTCTGCATATACAT 360
Qy 361 AACTGTATTCACAGAGATATAAAACCTGAAAATATTTCTAATAACTAAGCAAGGAATAATC 420
Db 361 AACTGTATTCACAGAGATATAAAACCTGAAAATATTTCTAATAACTAAGCAAGGAATAATC 420
Qy 421 AAGATTTGTGACTTCGGGTTTGCACAAATTTCTGATCCAGGAGATGCCATACACCGATTAT 480
Db 421 AAGATTTGTGACTTCGGGTTTGCACAAATTTCTGATCCAGGAGATGCCATACACCGATTAT 480
Qy 481 GTAGCTACGAGATGGTACCGAGCTCCTGAACCTTCTTGTGGGAGATACCTCAGTATGGTTCT 540
Db 481 GTAGCTACGAGATGGTACCGAGCTCCTGAACCTTCTTGTGGGAGATACCTCAGTATGGTTCT 540
Qy 541 TCAGTCGATATATGGGCTATTGGTTGTTTTCAGAGCTCCTGACAGGCCAGCCACTG 600
Db 541 TCAGTCGATATATGGGCTATTGGTTGTTTTCAGAGCTCCTGACAGGCCAGCCACTG 600

Db 541 TCAGTCGATATATGGGCTATTGGTTGTGTTTTTGCAGAGCTCCTGACAGGCCCACTG 600

Qy 601 TGGCCTGGAAATCAGATGAGGACCAACTTTTATCTGATAATCAGAACACTAGGAAATTA 660

Db 601 TGGCCTGGAAATCAGATGAGGACCAACTTTTATCTGATAATCAGAACACTAGGAAATTA 660

Qy 661 ATCCCAAGACATCAATCAATCTTTAAAGTAACGGGTTTTTCCATGGCATCAGTATACCT 720

Db 661 ATCCCAAGACATCAATCAATCTTTAAAGTAACGGGTTTTTCCATGGCATCAGTATACCT 720

Qy 721 GAGCCAGAAGACATGGAACCTCTTGAGGAAAGTTCTCAGATGTTTCATCCTGTGGCTCTG 780

Db 721 GAGCCAGAAGACATGGAACCTCTTGAGGAAAGTTCTCAGATGTTTCATCCTGTGGCTCTG 780

Qy 781 AACTTCATGAAGGGGTGCTGAAGATGAATCCAGATGACAGATTAACCTGTTCCTCAACTC 840

Db 781 AACTTCATGAAGGGGTGCTGAAGATGAATCCAGATGACAGATTAACCTGTTCCTCAACTC 840

Qy 841 CTGGAGAGCTCCTACTTTGATTTCTTTCAAGAGGCCCAAAATTAAGAAAGCAACGTAAT 900

Db 841 CTGGAGAGCTCCTACTTTGATTTCTTTCAAGAGGCCCAAAATTAAGAAAGCAACGTAAT 900

Qy 901 GAAGGAAGAAACAGAGACGCCCAACAGGT 929

Db 901 GAAGGAAGAAACAGAGACGCCCAACAGGT 929

RESULT 3

AX698820

LOCUS AX698820 1678 bp DNA linear PAT 02-APR-2003

DEFINITION Sequence 6 from Patent WO02079473.

ACCESSION AX698820

VERSION AX698820.1 GI:29499608

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE

AUTHORS Jones,A.L., Dam,T.C., Liu,T.F., Harris,B., Flores,V., Daffo,A., Marwaha,R., Chen,A.J., Chang,S.C., Gerstein,J.E., Peralta,C.H., David,M.H. and Lewis,S.A.

TITLE Molecules for diagnostics and therapeutics

JOURNAL Patent: WO 02079473-A 6 10-OCT-2002;

INCYTE Genomics, Inc. (US)

FEATURES

source

1. .1678

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/note="Incyte ID No: LI:058298.1:2001JAN12"

BASE COUNT 541 a 308 c 336 g 493 t

ORIGIN

Query Match 91.9%; Score 868.8; DB 6; Length 1678;

Best Local Similarity 99.2%; Pred. No. 2.2e-188;

Matches 925; Conservative 0; Mismatches 2; Indels 5; Gaps 5;

Qy 1 ATGGAAAAGTATGAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATCAAA 60

Db 131 ATGGAGAAGTATGAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATCAAA 190

Qy 61 TGCAGA-AACAAACCTCTGGACAAGTAGCTGTTAAATAATTTGTGGAATCTGAAGA 119

Db 191 TGCAGATAACAAACCTCTGGACAAGTAGCTGTTAAATAATTTGTGGAATCTGAAGA 250

Qy 120 TGATCCTGTTGTTAAGAAAATAGCACTAAGAGAAATACGTA-TGTTGAAGCAATTAAC 178

Db 251 TGATCCTGTTGTTAAGAAAATAGCACTAAGAGAAATACGTAAGCAATTAAC 310

Qy 179 ATCCAAATCTTGAAACCTTCATCGAGGTGTTTCAGGAGAAAAGCAAAATGCATTAGTTT 238

Db 311 ATCCAAATCTTGGAACCTCATCGAGGTGTTTCAGGAGAAAAGGAAAATGCATTAGTTT 370

Qy 239 TTGAATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTG 298

Db 371 TTGAATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTG 430

Qy 299 ATGGAGTGATCAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGTCAATATAC 358

Db 431 ATGGAGTGATCAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGTCAATATAC 490

Qy 359 ATAACTGTATTACAGAGATATATAAACCCTGAAAATATTCTAATAACTAAGCAAGGAATAA 418

Db 491 ATAACTGTATTACAGAGATATATAAACCCTGAAAATATTCTAATAACTAAGCAAGGAATAA 550

Qy 419 TCAAGATTTGTGACTTCGGGTTTGCACAAATTTCTGATTCAGGAGATGCCTACACCGATT 478

Db 551 TCAAGATTTGTGACTTCGGGTTTGCACAAATTTCTGATTCAGGAGATGCCTACACCGATT 610

Qy 479 ATGTAGCTA-CGAGATGGTACCAGCT-CCTGAACCTTCTTGTGGGAGATACT-CAGTATG 535

Db 611 ATGTAGCTAGCGAGATGGTACCAGCTCCTGAACTTCTTGTGGGAGATACTCCAGTATG 670

Qy 536 GTTCTTCAGTCGATATATGGGCTATTGGTTGTGTTTTTGCAGAGCTCCTGACAGGCCAGC 595

Db 671 GTTCTTCAGTCGATATATGGGCTATTGGTTGTGTTTTTGCAGAGCTCCTGACAGGCCAGC 730

Qy 596 CACTGTGGCCTGGAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGAA 655

Db 731 CACTGTGGCCTTGAATAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGAA 790

Qy 656 AATTAATCCCAAGACATCAATCAATCTTTAAAAAGTAAACGGTTTTTCCATGGCATCAGTA 715

Db 791 AATTAATCCCAAGACATCAATCAATCTTTAAAAAGTAAACGGTTTTTCCATGGCATCAGTA 850

Qy 716 TACCTGAGCCAGAGACATGGAACCTCTTGAGGAAAAGTTCTCAGATGTTTCATCCTGTGG 775

Db 851 TACCTGAGCCAGAGACATGGAACCTCTTGAGGAAAAGTTCTCAGATGTTTCATCCTGTGG 910

Qy 776 CTCTGAACCTTCATGAAGGGGTGCTGAAGATGAATCCAGATGACAGATTAACCTGTTCCC 835

Db 911 CTCTGAACCTTCATGAAGGGGTGCTGAAGATGAATCCAGATGACAGATTAACCTGTTCCC 970

Qy 936 AACTCCTGGAGAGCTCCTACTTTGATTCTTTCAAGAGGCCCAAAATTAAGAAAGAACAC 895

Db 971 AACTCCTGGAGAGCTCCTACTTTGATTCTTTCAAGAGGCCCAAAATTAAGAAAGAACAC 1030

Qy 896 GTAATGAAGGAAGAAACAGAGACGCCCAACAG 927

Db 1031 GTAATGAAGGAAGAAACAGAGACGCCCAACAG 1062

RESULT 4

AX166534

LOCUS AX166534 1083 bp DNA linear PAT 22-JUN-2001

DEFINITION Sequence 25 from Patent WO0138503.

ACCESSION AX166534

VERSION AX166534.1 GI:14546879

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE

AUTHORS Plowman,G.D., Whyte,D., Manning,G.S., Sudarsanam,S.S., Martinez,R., Flanagan,P. and Clary,D.S.

TITLE Novel human protein kinases and protein kinase-like enzymes

JOURNAL Patent: WO 0138503-A 25 31-MAY-2001;

SUGEN, Inc. (US)

FEATURES

source

1. .1083

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

BASE COUNT 366 a 199 c 225 g 293 t
ORIGIN

```
Query Match      91.3%; Score 863; DB 6; Length 1083;
Best Local Similarity 94.5%; Pred. No. 4.9e-187;
Matches 927; Conservative 0; Mismatches 0; Indels 54; Gaps 1;

QY 1 ATGGAAGATGATAAATAATAGCTAAGACTGGAGAAGGGTCTTATGGGTTGTATTCAA 60
   |||||||
Db 1 ATGGAAGATGATAAATAATAGCTAAGACTGGAGAAGGGTCTTATGGGTTGTATTCAA 60

QY 61 TGCAGAAACCAAAACCTCTGACCAAGTAGTAGCTGTTAAAAAATTTGTGGAATCTGAAGAT 120
   |||||||
Db 61 TGCAGAAACCAAAACCTCTGACCAAGTAGTAGCTGTTAAAAAATTTGTGGAATCTGAAGAT 120

QY 121 GATCCTGTTGTTAAGAAAATAGCACTAAGAGAAAATACGTATGTTGAAGCAATTAATAACAT 180
   |||||||
Db 121 GATCCTGTTGTTAAGAAAATAGCACTAAGAGAAAATACGTATGTTGAAGCAATTAATAACAT 180

QY 181 CCAATCTTGTGAACCTCATCGAGGTGTTTCAAGGAGAAAAGGAAAATGCATTTAGTTTTT 240
   |||||||
Db 181 CCAATCTTGTGAACCTCATCGAGGTGTTTCAAGGAGAAAAGGAAAATGCATTTAGTTTTT 240

QY 241 GAATACCTGATCATACACCTTTTAAATGAGCTGGAAAGAAAACCAAAATGGAGTTGCTGAT 300
   |||||||
Db 241 GAATACCTGATCATACACCTTTTAAATGAGCTGGAAAGAAAACCAAAATGGAGTTGCTGAT 300

QY 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCAATATACAT 360
   |||||||
Db 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCAATATACAT 360

QY 361 AACTGTATTACAGAGATATAAAACCTGAAAATATTTCTAATACTAAGCAAGGAATAATC 420
   |||||||
Db 361 AACTGTATTACAGAGATATAAAACCTGAAAATATTTCTAATACTAAGCAAGGAATAATC 420

QY 421 AAGATTTGTGACTTCGGGTTTGCACAAATTTCTGA----- 454
   |||||||
Db 421 AAGATTTGTGACTTCGGGTTTGCACAAATTTCTGAGTAATTCACATGTTGGGAGAAATTGAC 480

QY 455 -----TTCCAGGAGATGCCTACACCGATTATGTAGCT 486
   |||||||
Db 481 TTACCTGATCTGATTGATGCATTTGCAGTTCCAGGAGATGCCTACACCGATTATGTAGCT 540

QY 487 ACGAGATGTAACCGAGCTCCTGAACCTCTTGTGGAGATACTCAGTATGTTCTTCAGTC 546
   |||||||
Db 541 ACGAGATGTAACCGAGCTCCTGAACCTCTTGTGGAGATACTCAGTATGTTCTTCAGTC 600

QY 547 GATATATGGGTATTTGGTTGTGTTTTCAGAGACTCCTGACAGCCAGCCACTGTGGCCT 606
   |||||||
Db 601 GATATATGGGTATTTGGTTGTGTTTTCAGAGACTCCTGACAGCCAGCCACTGTGGCCT 660

QY 607 GGAATAATCAGATGTGGACCAACTTTTATCTGATAATCAGAAACACTAGGAAAATTAATCCCA 666
   |||||||
Db 661 GGAATAATCAGATGTGGACCAACTTTTATCTGATAATCAGAAACACTAGGAAAATTAATCCCA 720

QY 667 AGACATCAATCAATCTTTAAAGTAACGGGTTTTTCCATGGCATCAGTATACCTGAGCCA 726
   |||||||
Db 721 AGACATCAATCAATCTTTAAAGTAACGGGTTTTTCCATGGCATCAGTATACCTGAGCCA 780

QY 727 GAAGACATGAAAACCTCTTGAGGAAAAGTTCTCAGATGTTTCATCCTGTGGCTCTGAACCTC 786
   |||||||
Db 781 GAAGACATGAAAACCTCTTGAGGAAAAGTTCTCAGATGTTTCATCCTGTGGCTCTGAACCTC 840

QY 787 ATGAAGGGGTGCTGAAGATGAATCCAGATGACAGATTAACCTGTTCCCAACTCCTGGAG 846
   |||||||
Db 841 ATGAAGGGGTGCTGAAGATGAATCCAGATGACAGATTAACCTGTTCCCAACTCCTGGAG 900

QY 847 AGCTCCTACTTTGATTTCTTTCAAGAGGCCCCAAATTAAGAAAAGCACGTAATGAAGGA 906
   |||||||
Db 901 AGCTCCTACTTTGATTTCTTTCAAGAGGCCCCAAATTAAGAAAAGCACGTAATGAAGGA 960

QY 907 AGAAACAGAGAGCCCAACAG 927
   |||||||
```

Db 961 AGAAACAGAGACGCCCAACAG 981

RESULT 5
AX107720

LOCUS AX107720 1041 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 9 from Patent WO0123579.
ACCESSION AX107720
VERSION AX107720.1 GI:13923201
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 Donohc,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and
Sands,A.T.

HUMAN KINASE PROTEINS AND POLYNUCLEOTIDES ENCODING THE SAME

PATENT: WO 0123579-A 9 05-APR-2001;

LEXICON GENETICS INCORPORATED (US)

FEATURES
Location/Qualifiers

1..1041
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 344 a 190 c 228 g 279 t
ORIGIN

Query Match 88.8%; Score 839; DB 6; Length 1041;
Best Local Similarity 90.8%; Pred. No. 1.5e-181;
Matches 945; Conservative 0; Mismatches 0; Indels 96; Gaps 1;

QY 1 ATGGAAGATGAAAAATTAGCTAAGACTGGAGAAGGGTCTTATGGGTTGTATTCAA 60
|||

Db 1 ATGGAAGATGAAAAATTAGCTAAGACTGGAGAAGGGTCTTATGGGTTGTATTCAA 60
|||

QY 61 TGCAGAAACCAAAACCTCTGGCAAGTAGTAGCTGTTAAAAAATTTGTGGAATCTGAAGAT 120
|||

Db 61 TGCAGAAACCAAAACCTCTGGCAAGTAGTAGCTGTTAAAAAATTTGTGGAATCTGAAGAT 120
|||

QY 121 GATCCTGTTTAAAGAAAATAGCACTAAGAGAAAATACGTATGTTGAAGCAATTAATAACAT 180
|||

Db 121 GATCCTGTTTAAAGAAAATAGCACTAAGAGAAAATACGTATGTTGAAGCAATTAATAACAT 180
|||

QY 181 CCAATCTTGTGAACCTCATCGAGGTGTTCAAGGAGAAAAGGAAAATGCATTTAGTTTTT 240
|||

Db 181 CCAATCTTGTGAACCTCATCGAGGTGTTCAAGGAGAAAAGGAAAATGCATTTAGTTTTT 240
|||

QY 241 GAATACCTGTGATCATAACCTTTTAAATCAGCTGGAAAAGAACCCAAATGGAGTTGCTGAT 300
|||

Db 241 GAATACCTGTGATCATAACCTTTTAAATCAGCTGGAAAAGAACCCAAATGGAGTTGCTGAT 300
|||

QY 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCAATATACAT 360
|||

Db 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCAATATACAT 360
|||

QY 361 AACTGTATTCACAGAGATATAAAACCTGAAAATATTTCTAATACTAAGCAAGGAATAATC 420
|||

Db 361 AACTGTATTCACAGAGATATAAAACCTGAAAATATTTCTAATACTAAGCAAGGAATAATC 420
|||

QY 421 AAGATTTGTGACTTCGGGTTTGCACAAATTTCTGATTCAGGAGATGCCTACACCGATTAT 480
|||

Db 421 AAGATTTGTGACTTCGGGTTTGCACAAATTTCTGATTCAGGAGATGCCTACACCGATTAT 480
|||

QY 481 GTAGCTACGAGATGGTACCGAGCTCCTGAACTTCTTGTGGGAGATACTCAGTATGGTTCT 540
|||

Db 481 GTAGCTACGAGATGGTACCGAGCTCCTGAACTTCTTGTGGGAGATACTCAGTATGGTTCT 540
|||

QY 541 TCAGTCGATATATGGGCTATTGGTTGTGTTTTTGCAGAGCTCCTGACAGGCCAGCCACTG 600
|||

Db 541 TCAGTCGATATATGGGCTATTGGTTGTGTTTTTGCAGAGCTCCTGACAGGCCAGCCACTG 600
|||

QY 601 TGGCCTGGAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACT----- 650
|||

Db 601 TGCCCTGGAAATCAGATGGACCAACTTTATCTGATAATCAGAACACTAGTAGAGACG 660
Qy 651 ----- 650
Db 661 GGGTTTCGCCATGTTGACCCAGCGTGGTCTCGAACCTTTGACGTCAAGTGATCCACCTGCC 720
Qy 651 -----AGGAAATTAATCCCAAGACATCAATCAATCTTT 684
Db 721 GTAGCCTCTCAAAGTCTGGAATTACAGGAAATTAATCCCAAGACATCAATCAATCTTT 780
Qy 685 AAAAGTAACGGGTTTTTCCATGGCATCAGTATACCTGAGCCAGAGACATGGAAACTCTT 744
Db 781 AAAAGTAACGGGTTTTTCCATGGCATCAGTATACCTGAGCCAGAGACATGGAAACTCTT 840
Qy 745 GAGGAAAGTTCTCAGATGTTTCATCCTGTGGCTCTGAACTTCATGAAGGGGTGCTGAAG 804
Db 841 GAGGAAAGTTCTCAGATGTTTCATCCTGTGGCTCTGAACTTCATGAAGGGGTGCTGAAG 900
Qy 805 ATGAATCCAGATGACAGATTAACTGTTCCCAACTCTGAGAGAGCTCCTACTTTGATTCT 864
Db 901 ATGAATCCAGATGACAGATTAACTGTTCCCAACTCTGAGAGAGCTCCTACTTTGATTCT 960
Qy 865 TTTCAGAGGGCCCCAAATTAAAGAAAGACGTAATGAAGGAAGAAACAGAGACGCCAA 924
Db 961 TTTCAGAGGGCCCCAAATTAAAGAAAGACGTAATGAAGGAAGAAACAGAGACGCCAA 1020
Qy 925 CAGGTACTTCCGCTCAAAAGT 945
Db 1021 CAGGTACTTCCGCTCAAAAGT 1041
RESULT 6
AX107714
LOCUS AX107714 1068 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 3 from Patent WO0123579.
ACCESSION AX107714
VERSION AX107714.1 GI:13923198
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and Sands,A.T.
TITLE Human kinase proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0123579-A 3 05-APR-2001;
Lexicon Genetics Incorporated (US)
FEATURES
source
1..1068
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 344 a 197 c 240 g 287 t
ORIGIN
Query Match 87.1%; Score 823; DB 6; Length 1068;
Best Local Similarity 90.6%; Pred. No. 7e-178;
Matches 929; Conservative 0; Mismatches 0; Indels 96; Gaps 1;
Qy 1 ATGGAAAGTATGAAATAATTAGCTAAGACTGGAGAAGGGTCTTATGGGGTTGTATTCAA 60
Db 1 ATGGAAAGTATGAAATAATTAGCTAAGACTGGAGAAGGGTCTTATGGGGTTGTATTCAA 60
Qy 61 TGCAGAAACAAACCTCTGGACAAGTAGCTGTTAAAAAATTTGTGGAATCTGAAGAT 120
Db 61 TGCAGAAACAAACCTCTGGACAAGTAGCTGTTAAAAAATTTGTGGAATCTGAAGAT 120
Qy 121 GATCCTGTTGTTAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAAACAT 180
Db 121 GATCCTGTTGTTAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAAACAT 180

Qy 181 CCAAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAAGGAAATGCATTTAGTTTTT 240
Db 181 CCAAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAAGGAAATGCATTTAGTTTTT 240
Qy 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT 300
Db 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT 300
Qy 301 GGAGTGATCAAAGCGTATTATGGCAAACACTTCAAGCTCTTAAATTTCTGTATATACAT 360
Db 301 GGAGTGATCAAAGCGTATTATGGCAAACACTTCAAGCTCTTAAATTTCTGTATATACAT 360
Qy 361 AACTGTATTACAGAGATATAAAACCTGAAATAATTCTATAAATAAGCAAGGAATAATC 420
Db 361 AACTGTATTACAGAGATATAAAACCTGAAATAATTCTATAAATAAGCAAGGAATAATC 420
Qy 421 AAGATTTGTGACTTCGGGTTTGCACAAATTTCTGATTCAGGAGATGCCTACACCGATTAT 480
Db 421 AAGATTTGTGACTTCGGGTTTGCACAAATTTCTGATTCAGGAGATGCCTACACCGATTAT 480
Qy 481 GTAGCTACGAGATGGTACCGAGCTCCTGAACTTCTTGTGGAGATACCTCAGTATGGTTCT 540
Db 481 GTAGCTACGAGATGGTACCGAGCTCCTGAACTTCTTGTGGAGATACCTCAGTATGGTTCT 540
Qy 541 TCAGTCGATATATGGGCTATTGGTTGTGTTTTTGCAGAGCTCCTGACAGGCCAGCCACTG 600
Db 541 TCAGTCGATATATGGGCTATTGGTTGTGTTTTTGCAGAGCTCCTGACAGGCCAGCCACTG 600
Qy 601 TGGCCTGGAAATCAGATGTGGACCAACTTTATCTGATAATCAGAAACACT----- 650
Db 601 TGGCCTGGAAATCAGATGTGGACCAACTTTATCTGATAATCAGAAACACTTAGTAGAGACG 660
Qy 651 ----- 650
Db 661 GGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTCTTGCAGTCAAGTGATCCACCTGCC 720
Qy 651 -----AGGAAATTAATCCCAAGACATCAATCAATCTTT 684
Db 721 GTAGCCTCTCAAAGTCTGGAATTACAGGAAATTAATCCCAAGACATCAATCAATCTTT 780
Qy 685 AAAAGTAACGGGTTTTTCCATGGCATCAGTATACCTGAGCCAGAGACATGGAAACTCTT 744
Db 781 AAAAGTAACGGGTTTTTCCATGGCATCAGTATACCTGAGCCAGAGACATGGAAACTCTT 840
Qy 745 GAGGAAAGTTCTCAGATGTTTCATCCTGTGGCTCTGAACTTCATGAAGGGGTGCTGAAG 804
Db 841 GAGGAAAGTTCTCAGATGTTTCATCCTGTGGCTCTGAACTTCATGAAGGGGTGCTGAAG 900
Qy 805 ATGAATCCAGATGACAGATTAACTGTTCCCAACTCTGAGAGAGCTCCTACTTTGATTCT 864
Db 901 ATGAATCCAGATGACAGATTAACTGTTCCCAACTCTGAGAGAGCTCCTACTTTGATTCT 960
Qy 865 TTTCAGAGGGCCCCAAATTAAAGAAAGACGTAATGAAGGAAGAAACAGAGACGCCAA 924
Db 961 TTTCAGAGGGCCCCAAATTAAAGAAAGACGTAATGAAGGAAGAAACAGAGACGCCAA 1020
Qy 925 CAGGT 929
Db 1021 CAGGT 1025
RESULT 7
AX286069
LOCUS AX286069 687 bp DNA linear PAT 20-NOV-2001
DEFINITION Sequence 3 from Patent WO0179488.
ACCESSION AX286069
VERSION AX286069.1 GI:17045995
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1

AUTHORS Kapeller-Libermann,R.
TITLE 14257, protein kinase molecules and uses therefor
JOURNAL Patent: WO 0179488-A 3 25-OCT-2001;
Millennium Pharmaceuticals, Inc. (US)

FEATURES Location/Qualifiers

source
1..587
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 230 a 108 c 153 g 196 t

ORIGIN

Query Match 68.5%; Score 647; DB 6; Length 687;

Best Local Similarity 99.2%; Pred. No. 1.4e-137;

Matches 650; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ATGGAAAAGTATGAAAAATTAGCTAAGACTGGAGAAGGGCTTTATGGGGTTGTATTCAA 60

Db 1 ATGGAAAAGTATGAAAAATTAGCTAAGACTGGAGAAGGGCTTTATGGGGTTGTATTCAA 60

QY 61 TGCAGAAACAAACCTCTGGCAACAGTAGCTGTTAAAAAATTGTGGAATCTGAAGAT 120

Db 61 TGCAGAAACAAACCTCTGGCAACAGTAGCTGTTAAAAAATTGTGGAATCTGAAGAT 120

QY 121 GATCCTGTTTAAGAAAAATAGCACTAAGAGAAAAATACGTTATGTAAGCAATTAACAACAT 180

Db 121 GATCCTGTTTAAGAAAAATAGCACTAAGAGAAAAATACGTTATGTAAGCAATTAACAACAT 180

QY 181 CCAAACTCTGTGAACCTCATCGAGGTTTCAGGAGAAAAAGGAAAAATGCAATTTAGTTTT 240

Db 181 CCAAACTCTGTGAACCTCATCGAGGTTTCAGGAGAAAAAGGAAAAATGCAATTTAGTTTT 240

QY 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCCAATGGAGTTGCTGAT 300

Db 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCCAATGGAGTTGCTGAT 300

QY 301 GGAGTGATCAAAAGCGTATTATGGCAACACCTTCAAGCTCTTAAATTTCTGTCATATACAT 360

Db 301 GGAGTGATCAAAAGCGTATTATGGCAACACCTTCAAGCTCTTAAATTTCTGTCATATACAT 360

QY 361 AACTGTATTCACAGAGATATAAAACCTGAAAAATATTCTAATAACTAAGCAAGGAATAATC 420

Db 361 AACTGTATTCACAGAGATATAAAACCTGAAAAATATTCTAATAACTAAGCAAGGAATAATC 420

QY 421 AAGATTTGTGACTTCGGGTTTGCACAAATTCTGATTCAGGAGATGCCTACACCGATTAT 480

Db 421 AAGATTTGTGACTTCGGGTTTGCACAAATTCTGATTCAGGAGATGCCTACACCGATTAT 480

QY 481 GTAGCTACGAGATGGTACCGAGCTCCTGAACCTTCTTGGGAGATACTCAGTATGGTTCT 540

Db 481 GTAGCTACGAGATGGTACCGAGCTCCTGAACCTTCTTGGGAGATACTCAGTATGGTTCT 540

QY 541 TCAGTCGATATATGGGCTATTGGTTGTGTTTTTGCAGAGCTCCTGACAGGCCAGCCACTG 600

Db 541 TCAGTCGATATATGGGCTATTGGTTGTGTTTTTGCAGAGCTCCTGACAGGCCAGCCACTG 600

QY 601 TGGCCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGAA 655

Db 601 TGGCCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGTA 655

QY 601 TGGCCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGAA 655

Db 601 TGGCCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGTA 655

QY 601 TGGCCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGAA 655

Db 601 TGGCCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGTA 655

QY 601 TGGCCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGAA 655

Db 601 TGGCCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGTA 655

QY 601 TGGCCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGAA 655

Db 601 TGGCCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGTA 655

QY 601 TGGCCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGAA 655

Db 601 TGGCCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGTA 655

QY 601 TGGCCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGAA 655

Db 601 TGGCCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGTA 655

QY 601 TGGCCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGAA 655

Db 601 TGGCCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGTA 655

QY 601 TGGCCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGAA 655

Db 601 TGGCCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGTA 655

QY 601 TGGCCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGAA 655

Db 601 TGGCCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGTA 655

QY 601 TGGCCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGAA 655

Db 601 TGGCCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGTA 655

AUTHORS

Kapeller-Libermann,R.
TITLE 14257, protein kinase molecules and uses therefor
JOURNAL Patent: WO 0179488-A 1 25-OCT-2001;
Millennium Pharmaceuticals, Inc. (US)

FEATURES Location/Qualifiers

source
1..882
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 289 a 150 c 198 g 245 t

ORIGIN

Query Match 68.5%; Score 647; DB 6; Length 882;

Best Local Similarity 99.2%; Pred. No. 1.3e-137;

Matches 650; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ATGGAAAAGTATGAAAAATTAGCTAAGACTGGAGAAGGGCTTTATGGGGTTGTATTCAA 60

Db 1 ATGGAAAAGTATGAAAAATTAGCTAAGACTGGAGAAGGGCTTTATGGGGTTGTATTCAA 60

QY 61 TGCAGAAACAAACCTCTGGCAACAGTAGCTGTTAAAAAATTGTGGAATCTGAAGAT 120

Db 61 TGCAGAAACAAACCTCTGGCAACAGTAGCTGTTAAAAAATTGTGGAATCTGAAGAT 120

QY 189 TGCAGAAACAAACCTCTGGCAACAGTAGCTGTTAAAAAATTGTGGAATCTGAAGAT 248

Db 189 TGCAGAAACAAACCTCTGGCAACAGTAGCTGTTAAAAAATTGTGGAATCTGAAGAT 248

QY 121 GATCCTGTTTAAGAAAAATAGCACTAAGAGAAAAATACGTTATGTAAGCAATTAACAACAT 308

Db 121 GATCCTGTTTAAGAAAAATAGCACTAAGAGAAAAATACGTTATGTAAGCAATTAACAACAT 308

QY 181 CCAAACTCTGTGAACCTCATCGAGGTTTCAGGAGAAAAAGGAAAAATGCAATTTAGTTTT 240

Db 181 CCAAACTCTGTGAACCTCATCGAGGTTTCAGGAGAAAAAGGAAAAATGCAATTTAGTTTT 240

QY 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCCAATGGAGTTGCTGAT 300

Db 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCCAATGGAGTTGCTGAT 300

QY 369 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCCAATGGAGTTGCTGAT 428

Db 369 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCCAATGGAGTTGCTGAT 428

QY 301 GGAGTGATCAAAAGCGTATTATGGCAACACCTTCAAGCTCTTAAATTTCTGTCATATACAT 360

Db 301 GGAGTGATCAAAAGCGTATTATGGCAACACCTTCAAGCTCTTAAATTTCTGTCATATACAT 360

QY 429 GGAGTGATCAAAAGCGTATTATGGCAACACCTTCAAGCTCTTAAATTTCTGTCATATACAT 488

Db 429 GGAGTGATCAAAAGCGTATTATGGCAACACCTTCAAGCTCTTAAATTTCTGTCATATACAT 488

QY 361 AACTGTATTCACAGAGATATAAAACCTGAAAAATATTCTAATAACTAAGCAAGGAATAATC 420

Db 361 AACTGTATTCACAGAGATATAAAACCTGAAAAATATTCTAATAACTAAGCAAGGAATAATC 420

QY 489 AACTGTATTCACAGAGATATAAAACCTGAAAAATATTCTAATAACTAAGCAAGGAATAATC 548

Db 489 AACTGTATTCACAGAGATATAAAACCTGAAAAATATTCTAATAACTAAGCAAGGAATAATC 548

QY 421 AAGATTTGTGACTTCGGGTTTGCACAAATTCTGATTCAGGAGATGCCTACACCGATTAT 480

Db 421 AAGATTTGTGACTTCGGGTTTGCACAAATTCTGATTCAGGAGATGCCTACACCGATTAT 480

QY 549 AAGATTTGTGACTTCGGGTTTGCACAAATTCTGATTCAGGAGATGCCTACACCGATTAT 608

Db 549 AAGATTTGTGACTTCGGGTTTGCACAAATTCTGATTCAGGAGATGCCTACACCGATTAT 608

QY 481 GTAGCTACGAGATGGTACCGAGCTCCTGAACCTTCTTGGGGAGATACTCAGTATGGTTCT 540

Db 481 GTAGCTACGAGATGGTACCGAGCTCCTGAACCTTCTTGGGGAGATACTCAGTATGGTTCT 540

QY 609 GTAGCTACGAGATGGTACCGAGCTCCTGAACCTTCTTGGGGAGATACTCAGTATGGTTCT 668

Db 609 GTAGCTACGAGATGGTACCGAGCTCCTGAACCTTCTTGGGGAGATACTCAGTATGGTTCT 668

QY 541 TCAGTCGATATATGGGCTATTGGTTGTGTTTTTGCAGAGCTCCTGACAGGCCAGCCACTG 600

Db 541 TCAGTCGATATATGGGCTATTGGTTGTGTTTTTGCAGAGCTCCTGACAGGCCAGCCACTG 600

QY 669 TCAGTCGATATATGGGCTATTGGTTGTGTTTTTGCAGAGCTCCTGACAGGCCAGCCACTG 728

Db 669 TCAGTCGATATATGGGCTATTGGTTGTGTTTTTGCAGAGCTCCTGACAGGCCAGCCACTG 728

QY 601 TGGCCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGAA 655

Db 601 TGGCCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGTA 783

QY 601 TGGCCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGAA 655

Db 601 TGGCCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGTA 783

QY 601 TGGCCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGAA 655

Db 601 TGGCCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGTA 783

QY 601 TGGCCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGAA 655

Db 601 TGGCCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGTA 783

QY 601 TGGCCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGAA 655

Db 601 TGGCCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGTA 783

RESULT 9
AX056404

LOCUS AX056404 911 bp DNA linear PAT 13-JAN-2001
DEFINITION Sequence 48 from Patent WO0073469.
ACCESSION AX056404
VERSION AX056404.1 GI:12229111
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Plowman,G.D., Martinez,R., Whyte,D. and Sudersanam,S.
TITLE Protein kinases
JOURNAL Patent: WO 0073469-A 48 07-DEC-2000;
Sugen, Inc. (US)
FEATURES Location/Qualifiers
source
1..911
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 293 a 182 c 178 g 258 t
ORIGIN
Query Match 68.0%; Score 643; DB 6; Length 911;
Best Local Similarity 100.0%; Pred. No. 1.1e-136;
Matches 643; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 285 AAATGGAGTTGCTGATGGAGTGATCAAAAGCGTATTATGGCAACACCTCAAGCTCTTAA 344
Dd 1 AAATGGAGTTGCTGATGGAGTGATCAAAAGCGTATTATGGCAACACCTCAAGCTCTTAA 60
Qy 345 TTTCTGTCATATACATAACTGTTATTCACAGAGATATAAAACCTGAAATATTTCTAATAC 404
Dd 61 TTTCTGTCATATACATAACTGTTATTCACAGAGATATAAAACCTGAAATATTTCTAATAC 120
Qy 405 TAAGCAAGGAATAATCAAGATTTGTGACTTCGGGTTTGACAAATTCGATTCAGGAGA 464
Dd 121 TAAGCAAGGAATAATCAAGATTTGTGACTTCGGGTTTGACAAATTCGATTCAGGAGA 180
Qy 465 TGCCTACACCGATTATGTAGTACGAGATGGTACCGAGCTCCTGAACCTTCTGTGGGAGA 524
Dd 181 TGCCTACACCGATTATGTAGTACGAGATGGTACCGAGCTCCTGAACCTTCTGTGGGAGA 240
Qy 525 TACTCAGTATGGTTCTTCAGTCGATATATGGGCTATTGGTTGTGTTTTGGCAGAGCTCCT 584
Dd 241 TACTCAGTATGGTTCTTCAGTCGATATATGGGCTATTGGTTGTGTTTTGGCAGAGCTCCT 300
Qy 585 GACAGCCAGCCACTGTGGCCTGGAAATCAGATGTGGACCAACTTTATCTGATAATCAG 644
Dd 301 GACAGCCAGCCACTGTGGCCTGGAAATCAGATGTGGACCAACTTTATCTGATAATCAG 360
Qy 645 AACACTAGGAAATTAATCCCAAGACATCAATCAATCTTTAAAGTAACGGGTTTTTCCA 704
Dd 361 AACACTAGGAAATTAATCCCAAGACATCAATCAATCTTTAAAGTAACGGGTTTTTCCA 420
Qy 705 TGGCATCAGTATACCTGAGCCAGAGACATGGAAACTCTTGAGGAAAAGTTCTCAGATGT 764
Dd 421 TGGCATCAGTATACCTGAGCCAGAGACATGGAAACTCTTGAGGAAAAGTTCTCAGATGT 480
Qy 765 TCATCCTGTGGCTCTGAACTTCATGAAGGGGTGTCTGAAGATGAATCCAGATGACAGATT 824
Dd 481 TCATCCTGTGGCTCTGAACTTCATGAAGGGGTGTCTGAAGATGAATCCAGATGACAGATT 540
Qy 825 AACCTGTTCCCAACTCCTGGAGAGCTCCTACTTTGATCTTTTCAAGAGGCCCAATTA 884
Dd 541 AACCTGTTCCCAACTCCTGGAGAGCTCCTACTTTGATCTTTTCAAGAGGCCCAATTA 600
Qy 885 AAGAAAAGCAGCTAATGAAGGAAGAAACAGAGACGCGCAACAG 927
Dd 601 AAGAAAAGCAGCTAATGAAGGAAGAAACAGAGACGCGCAACAG 643

RESULT 10
AX056405

LOCUS AX056405 2615 bp DNA linear PAT 13-JAN-2001
DEFINITION Sequence 49 from Patent WO0073469.
ACCESSION AX056405
VERSION AX056405.1 GI:12229112
KEYWORDS
SOURCE Murinae gen. sp.
ORGANISM Murinae gen. sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae.
REFERENCE 1
AUTHORS Plowman,G.D., Martinez,R., Whyte,D. and Sudersanam,S.
TITLE Protein kinases
JOURNAL Patent: WO 0073469-A 49 07-DEC-2000;
Sugen, Inc. (US)
FEATURES Location/Qualifiers
source
1..2615
/organism="Murinae gen. sp."
/mol_type="genomic DNA"
/db_xref="taxon:39108"
BASE COUNT 742 a 580 c 585 g 708 t
ORIGIN
Query Match 59.2%; Score 559.2; DB 6; Length 2615;
Best Local Similarity 81.8%; Pred. No. 1.5e-117;
Matches 658; Conservative 0; Mismatches 143; Indels 3; Gaps 1;
Qy 135 GAAAATAGCACTAAGAGAAAATACGTATGTTGAAGCAATTAAACATCCAAATCTGTGAA 194
Dd 1 GAAAATAGCCCTGCGGGAATCCGTATGCTGAAG--TTGAAACACCCAAACCTCGTGAA 57
Qy 195 CCTCATCGAGGTGTTGAGGAGAAAAGGAAAATGCAATTTAGTTTTGAATACTGTGATCA 254
Dd 58 CCTCATCGAGGTGTTGAGGAGAAAAGAGAAAAGATGCATCTAGTTTTGAGTACTGTGATCA 117
Qy 255 TACACTTTTAAATGAGCTTGAAGAAAACCCAAATGGAGTTGCTGATGGAGTGATCAAAAG 314
Dd 118 CACACTGTTAAACGAGCTGGAGAGAAAACCCAAACGGAGTTTCTGATGGAGTGATTAAAG 177
Qy 315 CGTATTATGCCAAACACTTCAAGCTCTTAATTTCTGTCAATATACATAACTGTATTACAG 374
Dd 178 TGTGCTATGGCAAAACCTTCAAGCCCTTAACCTTCTGTCAACAGCACAATTGTATTATCG 237
Qy 375 AGATATAAAACCTGAAAATATTCTTAATAACTAAGCAAGGAATAATCAAGATTTGTGACTT 434
Dd 238 GGATGTAAAACCTGAAAACATCCTTAATAACCAAGCAAGGGATGATAAAGATTTGTGACTT 297
Qy 435 CGGGTTTGCACAAATTTCTGATTCAGGAGATGCCTACACCGGATTTATGTAGCTACGAGATG 494
Dd 298 TGGATTTGCACGAATTTCTAATTTCCAGGAGACGCCTACACAGACTATGTTGCCACCAGTG 357
Qy 495 GTACCGAGCTCCTGAACCTTCTTGTGGGAGATACACTCAGTATGGTTCTTCAGTCGATATATG 554
Dd 358 GTACCGAGCCCCGAACTTCTCGTGGGAGACACGAAGTACGGTTCTCTGTAGACGTGTG 417
Qy 555 GGCTATTGGTTGTGTTTTTGCAGAGCTCCTGACAGGCCAGCCACTGTGGCCTGGAAAATC 614
Dd 418 GGCCGTGCGCTGTGTTTTTGCAGAGCTCCTGACGGGTGAGCCACTCTGGCCGGGAAAATC 477
Qy 615 AGATGTGGACCAACTTTATCTGATAATCAGAAACACTAGGAAAATAATATCCCAAGACATCA 674
Dd 478 CGACGTGGACCACTTTACCTGATCATCAGGACGTTGGGGAAGCTGATTCCAAGACACCA 537
Qy 675 ATCAATCTTTAAAAGTAACGGGTTTTTCCATGGCATCAGTATACCTGAGCCAGAGACAT 734
Dd 538 GTCTATCTTTAGGAGTAACCAAGTTTTTCCGCGGCATCAGCATACCTGAACCCAGAGGACAT 597
Qy 735 GGAAACTCTTGAGAAAAGTTCTCAGATGTTTCATCTGTGGCTCTGAACCTTCATGAAGGG 794
Dd 598 GGAGACTCTTTGAAGAAAATTTCTCAAAATGTTTCAGCCTGTGGCTTTAAGTTTTCATGAAGGG 657
Qy 795 GTGTCTGAAGATGAATCCAGATGACAGATTAACCTGTTCCCAACTCCTGGAGAGCTCCTA 854
Dd 658 ATGCCTGAAGATGAATCCTGATGAGAGGCTGACCTGTGCCCCAGCTGCTGGACAGTGCCCTA 717

TITLE Human kinase proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0123579-A 1 05-APR-2001;
Lexicon Genetics Incorporated (US)
FEATURES
source 1. .561
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 189 a 87 c 112 g 173 t
ORIGIN
Query Match 48.1%; Score 454.4; DB 6; Length 561;
Best Local Similarity 99.8%; Pred. No. 1.7e-93;
Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGAAGAAGTATGAAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGGTTGTATTCAAA 60
DB 1 ATGGAAGAAGTATGAAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGGTTGTATTCAAA 60
QY 61 TGCAGAAACAAACCTCTGACAAAGTAGTAGCTGTTAAAAAATTTGTGGAATCTGAAGAT 120
DB 61 TGCAGAAACAAACCTCTGACAAAGTAGTAGCTGTTAAAAAATTTGTGGAATCTGAAGAT 120
QY 121 GATCCTGTTGTTAAGAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAAAAACAT 180
DB 121 GATCCTGTTGTTAAGAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAAAAACAT 180
QY 181 CCAATCTTGTGAACCTCATCGAGGTGTTCCAGGAGAAAAAGGAAAATGCAATTTAGTTTTT 240
DB 181 CCAATCTTGTGAACCTCATCGAGGTGTTCCAGGAGAAAAAGGAAAATGCAATTTAGTTTTT 240
QY 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATCGAGTTGCTGAT 300
DB 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATCGAGTTGCTGAT 300
QY 301 GGAGTGATCAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGTCATATACAT 360
DB 301 GGAGTGATCAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGTCATATACAT 360
QY 361 AACTGTATTCACAGAGATATAAAACCTGAAATATTCTAATAACTAAGCAAGGAATAATC 420
DB 361 AACTGTATTCACAGAGATATAAAACCTGAAATATTCTAATAACTAAGCAAGGAATAATC 420
QY 421 AAGATTGTGACTTCGGGTTTGCACAAATTTCTGATT 456
DB 421 AAGATTGTGACTTCGGGTTTGCACAAATTTCTGAGT 456
RESULT 13
AX107718
LOCUS AX107718 594 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 7 from Patent WO0123579.
ACCESSION AX107718
VERSION AX107718.1 GI:13923200
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1
AUTHORS Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and Sands,A.T.
TITLE Human kinase proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0123579-A 7 05-APR-2001;
Lexicon Genetics Incorporated (US)
FEATURES
source 1. .594
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 203 a 94 c 115 g 182 t
ORIGIN

Query Match 48.1%; Score 454.4; DB 6; Length 594;
Best Local Similarity 99.8%; Pred. No. 1.7e-93;
Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGAAGAAGTATGAAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGGTTGTATTCAAA 60
DB 1 ATGGAAGAAGTATGAAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGGTTGTATTCAAA 60
QY 61 TGCAGAAACAAACCTCTGACAAAGTAGTAGCTGTTAAAAAATTTGTGGAATCTGAAGAT 120
DB 61 TGCAGAAACAAACCTCTGACAAAGTAGTAGCTGTTAAAAAATTTGTGGAATCTGAAGAT 120
QY 121 GATCCTGTTGTTAAGAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAAAAACAT 180
DB 121 GATCCTGTTGTTAAGAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAAAAACAT 180
QY 181 CCAATCTTGTGAACCTCATCGAGGTGTTCCAGGAGAAAAAGGAAAATGCAATTTAGTTTTT 240
DB 181 CCAATCTTGTGAACCTCATCGAGGTGTTCCAGGAGAAAAAGGAAAATGCAATTTAGTTTTT 240
QY 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATCGAGTTGCTGAT 300
DB 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATCGAGTTGCTGAT 300
QY 301 GGAGTGATCAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGTCATATACAT 360
DB 301 GGAGTGATCAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGTCATATACAT 360
QY 361 AACTGTATTCACAGAGATATAAAACCTGAAATATTCTAATAACTAAGCAAGGAATAATC 420
DB 361 AACTGTATTCACAGAGATATAAAACCTGAAATATTCTAATAACTAAGCAAGGAATAATC 420
QY 421 AAGATTGTGACTTCGGGTTTGCACAAATTTCTGATT 456
DB 421 AAGATTGTGACTTCGGGTTTGCACAAATTTCTGAGT 456
RESULT 14
AX399975
LOCUS AX399975 1612 bp DNA linear PAT 06-JUN-2002
DEFINITION Sequence 146 from Patent WO0218424.
ACCESSION AX399975
VERSION AX399975.1 GI:21336301
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1
AUTHORS Tang,Y.T., Asundi,V., Zhou,P., Xue,A.J., Ren,F., Zhang,J., Wang,J.R., Zhao,Q.A., Wang,D., Liu,C., Drmanac,R.T. and Wehrman,T.
TITLE Nucleic acids and polypeptides
JOURNAL Patent: WO 0218424-A 146 07-MAR-2002;
HYSEQ, INC. (US)
FEATURES
Location/Qualifiers
source 1. .1612
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
463. .1539
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD33419.1"
/db_xref="GI:21336302"
/translation="MMKEYEKIGKIGESGYVVFVKCRNRTDGTQIIVAIKKFLESEDDPV
IKKIALREIRMLKQLKHPNLVNLLEVFRRKRLHLVFEYCDHTVLHLDYRQGVPEH
LVKSITWQTLQAVNFCHKNCIHRDVKPNILITKHSVIKLCDFGFARLLTGPSDYIT
DYVATRWYRSPPELLVGTQYQPPDVVAIGCVFAELLSGVPKDWPGKSDVDQLYLIRKT
LGDLPYRHOQVFTNQYFSGVKIPDPEDMEPLELKFNPISYPALGLKGLKGLHMDPTER
LTCEQLLHPYFENIEIDLAKENKPTKTRKSRKHKCFETSKLQYLPQLTGSS
ILPALDNKKYYCDTKLNYRFPNI"
BASE COUNT 476 a 380 c 361 g 395 t
ORIGIN

Query Match		44.0%;	Score 415.4;	DB 6;	Length 1612;
Best Local Similarity		66.7%;	Pred. No. 1.3e-84;		
Matches 609;		Conservative 0;	Mismatches 301;	Indels 3;	Gaps 1;
QY	1	ATGGAAAAGTATGAAAAAATTAGCTAAGACTGGAGAAGGGTCTTATGGGGTTGTATTCAAA	60		
DB	466	ATGGAGAAGTATGAAAAAATTGGGAAAATTGGAGAAGGATCCTATGGAGTTGTTTCAAA	525		
QY	61	TGCAGAAACAACCTCTCGACAAGTAGTAGCTGTTAAAAAATTTGTGGAATCTGAAGAT	120		
DB	526	TGTAGAAACAGGGACACGGGTACAGATTGTGGCCATCAAGAAGTTTCTTGGAAATCAGAAGAT	585		
QY	121	GATCCTGTTGTTAAGAAAATAGCACTAAGAGAAAATACGTATGTTGAAGCAATTAAAAACAT	180		
DB	586	GACCCCTGTCAAAAGAAAATTGCCCTTCGGGAAAATCCGAATGCTCAAGCAACTCAAGCAT	645		
QY	181	CCAAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAAGGAAAATGCAATTAGTTTTT	240		
DB	646	CCCAACCTTGTTAACTCTCTGGAAGTCTTCAGGAGGAAACGGAGGCTTCACCTGGTGT	705		
QY	241	GAATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT	300		
DB	706	GAATATTGTGACCAACACAGTTCTCCATGAGTTGGACAGATACCAAGAGGGGTACCAGAA	765		
QY	301	GGAGTGATCAAAAGCGTATTATGGCAACACACTCAAGCTCTTAATTTCTGTCTATATACAT	360		
DB	766	CATCTCGTGAAGACATAAATTGGCAGACACTGCAAGCTGTAATTTTGGCCATAAACAC	825		
QY	361	AACTGTATTACAGAGATATAAAACCTGAAAATATTCTAATAACTAAGCAAGGAATATC	420		
DB	826	AATTGCATACATAGAGACGTGAAGCCAGAAAATATCCTCATCAGAAACATTCCTGTGAT	885		
QY	421	AAGATTTGTGACTTCGGGTTTGCACAAAATTTCTGATT--CCAGGAGATGCCCTACACCGAT	477		
DB	886	AAGCTTTGTGACTTTGGATTTGCTCGGCTTTGACTGGACCGAGTGACTACTATACAGAC	945		
QY	478	TATGTAGCTACGAGATGGTACCGAGCTCCTGAACTTCTTGTTGGGAGATACCTCAGTATGGI	537		
DB	946	TACGTGGCTACCAGTGGTACCGCTCCCTGAGCTGCTGTTGGGGACACGCAGTACGGC	1005		
QY	538	TCTTCAGTCGATATATGGGCTATTGGTTGTGTTTTTGCAGAGCTCCTGACAGGCCAGCCA	597		
DB	1006	CCCCCGGTGATGTTTGGGCAATTGGCTGTGTCTTTGCTGAGCTGTGTACGGAGTGCCT	1065		
QY	598	CTGTGGCCTGGAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGAAAA	657		
DB	1066	CTGTGGCCAGGAAAATCGGATGTGGATCAGCTGTATCTGATTAGGAAGACCTTGGGGGAT	1125		
QY	658	TTAATCCCAAGACATCAATCAATCTTTAAAAAGTAACGGGTTTTTCCATGGCATCAGTATA	717		
DB	1126	CTCATTCCTAGGCACCAAGAGTGTGTTAGCAGCAATCAGTACTTCAGTGGAGTGAAAAAT	1185		
QY	718	CCTGAGCCAGAAAGACATGGAACCTCTTGAGGAAAAGTTCTCAGATGTTTCATCCTGTGGCT	777		
DB	1186	CCAGACCTTGAAGATATGGAACCACTTGAAATTAATAATTTCCCAACATCTCTTATCCTGCC	1245		
QY	778	CTGAACCTTCATGAAGGGGTGTCTGAAGATGAATCCAGATGACAGATTAACCTGTTCCCAA	837		
DB	1246	CTGGGGCTCCTAAAGGGCTGTCTCCACATGGACCCCTACTGAAAGGCTGACATGTGAACAG	1305		
QY	838	CTCCTGGAGAGCTCCTACTTTTGGATTCTTTTCAAGAGGCCCAAAATTAAGAAAGACACGT	897		
DB	1306	CTGTTGCATCACCCATATTTTGAAAACATCAGAGAAAATAGAGGATTTGGCAAAAGAACAC	1365		
QY	898	AATGAAGGAAGAA 910			
DB	1366	AACAACCAACAA 1378			

RESULT 15
HSSTHPKB
LOCUS HSSTHPKB 1363 bp mRNA linear PRI 06-FEB-1997

DEFINITION	H.sapiens mRNA KKIALRE for serine/threonine protein kinase.				
ACCESSION	X66358				
VERSION	X66358.1 GI:36614				
KEYWORDS	serine/threonine protein kinase.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 1363)				
AUTHORS	Meyerson,M., Enders,G.H., Wu,C.L., Su,L.K., Gorka,C., Nelson,C., Harlow,E. and Tsai,L.H.				
TITLE	A family of human cdc2-related protein kinases				
JOURNAL	EMBO J. 11 (8), 2909-2917 (1992)				
MEDLINE	92347325				
PUBMED	1639063				
REFERENCE	2 (bases 1 to 1363)				
AUTHORS	Meyerson,M.L.				
TITLE	Direct Submission				
JOURNAL	Submitted (12-MAY-1992) M.L. Meyerson, Massachusetts General Hospital, Cancer Center, Bldg 149, 13th Street, Charlestown MA 02129, USA				
FEATURES	Location/Qualifiers				
source	1..1363				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/cell_line="HeLa"				
	/cell_type="HeLa"				
	/clone_lib="HeLa"				
	214..1290				
	/codon_start=1				
	/product="serine/threonine protein kinase"				
	/protein_id="CAA47002.1"				
	/db_xref="GI:36615"				
	/db_xref="SWISS-PROT:Q00532"				
	/translation="MMEKYEKIGKIGESYGVVFKCRNRTGOIVAICKFLESEDDPV IKKIALREIRMLKQLKHPNLVLEVFRRKRLHLVFYCDHTVLHEDRYQGVPEH LVKSTTWOTLOAVNFCVHNCIHRDVKPENILITKHSVIKLCDFGFARLLTGPSDYVT DYVATRWYRSPPELLVGTQYQPPVDVWAIGCVFAELLSGVPLWPKSDVDQLYLIRKT LGDLIPRHOQVFSTNQYFSGVKIPDPEDMEPLEKFPNISYPALGLLKGCLHMDPTER LTCEQLJHPYFENIREIDLAKEHDKPTRKTLRKRKHCFTETSKLOYLPQLTGSS ILPALDNKKYVCDTKLNVFPNI"				
	join(214..669,X66359.1:1..54,670..1290)				
	/note="KKIALRE coding sequence with 54 bp insertion"				
	/pseudo				
	/codon_start=1				
	/label=schpkin_CDS				
	/product="serine/threonine protein kinase"				
	/db_xref="PID:e49306"				
BASE COUNT	412 a	308 c	309 g	334 t	
ORIGIN					
Query Match	43.8%; Score 413.8; DB 9; Length 1363;				
Best Local Similarity	66.6%; Pred. No. 3e-84;				
Matches	608; Conservative 0; Mismatches 302; Indels 3; Gaps 1;				
QY	1	ATGGAAAAGTATGAAAAATTAGCTAAGACTGGAGAAGGTCCTTATGGGTTGTATTCAA	60		
Db	217	ATGGAGAAGTATGAAAAAATTGGGAAAATTGGAGAAGGATCCTATGGAGTTGTTTCAA	276		
QY	61	TGCAGAAAACAAACCTCTCTGGACAAGTAGTAGCTGTTAAAAAATTTGTGGAATCTGAAGAT	120		
Db	277	TGTAGAAACAGGGACACCGGTTCAGATTGTGGCCATCAAGAAGTTTCTCGAATCAGAAGAT	336		
QY	121	GATCCTGTTGTTAAGAAAATAGCACTAAGAGAAAATACGTATGTTGAAGCAATTAACAACAT	180		
Db	337	GACCCCTGTCAATAAGAAAAATTGCCCTTCGGGAAAATCCGAATGCTCAAGCAACTCAAGCAT	396		
QY	181	CCAAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAGGAAAATGCAATTAGTTTTT	240		
Db	397	CCCAACCTTGTTAACTCTCTCGAAGTCTTCAGGAGGAAAACGGAGGCTTCACCTGGTGT	456		
QY	241	GAATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCCAATGGAGTTGCTGAT	300		

```

Db      ||||| ||||| || ||| || | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
457 GAATATTGTGACCACACAGTTCTCCATGAGTTGGACAGATACCCAAGAGGGGTACCAGAA 516

Qy      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
301 GGAGTGATCAAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAAATTTCTGTATATACAT 360

Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
517 CATCTCGTGAAGAGCATAAAGTGGCAGACACTGCAAGCTGTAAATTTTGCCATAAACAC 576

Qy      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
361 AACTGTATTACAGAGATATAAAACCTGAAAATATTCTAATAACTAAGCAAGGAATAATC 420

Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
577 AATTGCATACATAGAGACGTGAAGCCAGAAAATATCCTCATCAGAAACATTCGGTGATT 636

Qy      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
421 AAGATTGTGACITTCGGTTTGCACAAATTCGATT---CCAGGAGATGCCCTACACCGAT 477

Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
637 AAGCTTTGTGACITTTGGATTGTCTCGGCTTTTGACTGGACCGAGTGAATCTACTATACAGAC 696

Qy      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
478 TATGTAGCTACGAGATGGTACCAGAGCTCCTGAACTTCTTGIGGGAGATCTCAGTATGGT 537

Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
697 TACGTGGCTACCCAGGTGGTACCGCTCCCTGAGCTGCTGGTGGGGACACGCAGTACGGC 756

Qy      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
538 TCTTCAGTCGATATATGGGCTATTGGTTGTGTTTTCAGAGCTCCTGACAGGCCCGCCA 597

Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
757 CCCCCGGTGGATGTTTGGGCAATTGGCTGTGCTTGTGAGCTGCTGTCAGGAGTGCCT 816

Qy      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
598 CTGTGGCCTGGAAAATCAGATGTGGACCAACTTTATCTGATTAATCAGAAACACTAGGAAA 657

Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
817 CTGTGGCCAGGAAAATCGGATGTGGATCAGCTGTATCTGATTAGGAAGACCTTGGGGGAT 876

Qy      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
658 TTAATCCCAAGACATCAATCAATCTTTAAAGTAAAGGGTTTTCCATGGCATCAGTATA 717

Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
877 CTCATTCCTAGGCCACCGCAAGTGTTTAGCACGAATCAGTACTTCAGTGGAGTGAAAATT 936

Qy      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
718 CCTGAGCCAGAGACATGGAAACTCTTGAGGAAAAGTTCTCAGATGTTTCATCCTGTGGCT 777

Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
937 CCAGACCCCTGAAGATATGGAAACCACCTTGAATTAAAATTCCTCAACATCTCTTATCCTGCC 996

Qy      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
778 CTGAACCTTCATGAAGGGGTGTCTGAAGATGAATCCAGATGACAGATTAACTGTTCCTCAA 837

Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
997 CTGGGGCTCCTAAAGGGGTGTCTCCACATGGACCCCTACTGAAAGGGCTGACATGTGAACAG 1056

Qy      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
838 CTCTGGAGAGCTCCTACTTTTGATTTCTTTTCAAGAGGCCCAATTTAAAGAAAAGCACGT 897

Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1057 CTGTTGCATCACCCATATTTTGAAAACATCAGAGAAAATAGAGGATTTGGCAAAAGAACAC 1116

Qy      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
898 AATGAAGGAAGAA 910

Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1117 GACAAACCAACAA :129
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen, Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2003, 11:38:36 ; Search time 296.918 Seconds
(without alignments)
8591.48; Million cell updates/sec

Title: US-09-671-050-11

Perfect score: 945

Sequence: 1 atggaaaagtatgaaaaatt.....aggtacttcggtcaaaagt 945

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_19Jun03.*

- 1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
- 2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
- 3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
- 4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
- 5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
- 6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
- 7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
- 8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
- 9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
- 10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
- 11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
- 12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
- 13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
- 14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
- 15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
- 16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
- 17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
- 18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
- 19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
- 20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
- 21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
- 22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
- 23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
- 24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
- 25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	945	100.0	945	22	AAD03817 Human kinase cDNA
2	929	98.3	972	22	AAD03814 Human kinase cDNA
3	927	98.1	1790	24	AAD30557 Human kinase polyp
4	868.8	91.9	1678	25	ABX08936 cDNA encoding huma
5	863	91.3	1083	22	AAS06725 Polynucleotide seq
6	839	88.8	1041	22	AAD03816 Human kinase cDNA
7	823	87.1	1068	22	AAD03813 Human kinase cDNA
8	645.4	68.3	882	24	AAI64248 Human kinase 14257

9	643	68.0	911	22	AAF44669 Novel protein kina
10	559.2	59.2	2615	22	AAF44670 Novel protein kina
11	543.2	57.5	1281	25	ABX34679 Human mddt cDNA SE
12	507	53.7	1819	22	AAD03818 Human kinase cDNA
13	501	53.0	1086	25	ABZ77165 Human protein kina
14	454.4	48.1	561	22	AAD03812 Human kinase cDNA
15	454.4	48.1	594	22	AAD03815 Human kinase cDNA
16	415.4	44.0	1612	24	ABQ33433 Human cDNA SEQ ID
17	413.8	43.8	2944	23	ABV24584 Human prostate exp
18	287	30.4	1698	21	AAZ29746 Rabbit KKIAMRE kin
19	287	30.4	3080	21	AAZ29745 Rabbit KKIAMRE kin
20	274	29.0	1179	23	ABL12603 Drosophila melanog
21	227.4	24.1	1513	21	AAZ51208 Human lost in leuk
22	227.4	24.1	1773	22	AAS06724 Polynucleotide seq
23	219.4	23.1	3297	23	ABL12602 Drosophila melanog
24	211.4	22.4	1667	21	AAZ51207 Rat lost in leukae
25	211.4	22.4	1738	21	AAZ51206 DNA encoding novel
26	177.4	18.8	1158	23	AAS70329 Novel protein kina
27	175.8	18.6	1866	22	AAF44672 Pneumocystis carin
28	171.8	18.2	903	20	AAZ07476 Cyclin-dependent k
29	163.4	17.3	1476	21	AAZ37835 CDK2-green fluore
30	163.4	17.3	1635	19	AAV71073 Green fluorescent
31	163.4	17.3	1635	19	AAV71074 Toxicologically re
32	163	17.2	1024	25	ABZ83372 Human 10ckshs1 cDN
33	163	17.2	1297	24	ABQ60787 Human cDNA differe
34	161.8	17.1	2213	24	ABK84073 Human prostate can
35	161.4	17.1	1322	21	AAF15889 Human breast cance
36	155.6	16.5	1050	24	ABT11085 Lung cancer relate
37	155.6	16.5	1050	24	ABU64774 Lung cancer relate
38	155.6	16.5	1050	24	ABU65408 Lung cancer relate
39	155.6	16.5	1050	24	ABU65855 Lung cancer relate
40	155.6	16.5	1050	24	ABU69118 Kidney cancer rela
41	155.6	16.5	1050	25	ACC50096 Breast cancer asso
42	155.6	16.5	1050	25	AAD52783 Human CDC2 DNA. H
43	155.6	16.5	1235	25	ABX76248 Lung cancer-associ
44	155.6	16.5	1825	24	AAS94983 Human DNA sequence
45	155.6	16.5	1897	25	ACA03966 cDNA downregulated

ALIGNMENTS

RESULT 1

AAD03817 ID AAD03817 standard; cDNA; 945 BP.

XX AAD03817;

AC AAD03817;

XX 19-JUN-2001 (first entry)

DT Human kinase cDNA #6.

XX Human; kinase; gene therapy; bioreactor; mental disorder;

KW biological disorder; ss.

XX Homo sapiens.

OS Key

XX Location/Qualifiers

FT CDS

FT 1..945

FT /*tag= a

FT /product= "Human kinase #6"

FT /note= "The coding region does not include stop codon"

FT /partial

XX WO200123579-A1.

XX 05-APR-2001.

XX 27-SEP-2000; 2000WO-US26621.

XX 28-SEP-1999; 99US-0156511.

XX (LEXI-) LEXICON GENETICS INC.

PA

XX Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
PI WP1; 2001-266166/27.
XX P-PSDB; AAE00495.
DR
XX New isolated human kinase polynucleotide useful for generating
PT antibodies, as reagents in diagnostic assays and for screening for
PT compounds useful for treating mental, biological or medical diseases -
XX Claim 3; Page 33; 38pp; English.
PS
XX The present sequence is a cDNA encoding novel human protein
CC (NHP) known as human kinase. The human kinases share structural
CC similarity with animal kinases, more particularly serine or
CC threonine protein kinases. Human kinase cDNA is useful for the
CC detection of mutant human kinase for the diagnosis of disease,
CC and also as a therapeutic. It is useful for screening drugs
CC effective in the treatment of symptomatic or phenotypic
CC manifestations perturbing the normal function of NHP in the
CC body. The NHP nucleotide sequences are useful for generation of
CC antibodies, as reagents in diagnostic assays, for the
CC identification of other cellular gene products related to human
CC kinases, and as reagents in assays for screening compounds that
CC are useful for treating mental, biological or medical disorders.
CC NHP oligonucleotides are used as probes. The labelled NHP probes
CC are useful for screening human genomic library for identifying
CC polymorphisms and as primers in amplification assays to detect
CC mutations within the exons, introns and splice sites that can
CC be used in diagnostics and pharmacogenomics. Nucleotide construct
CC encoding NHP products are used to genetically engineer cells
CC in vivo that functions as bioreactors in the body delivering a
CC continuous supply of NHP to the body. Nucleotide constructs
CC encoding functional NHPs are used in gene therapy for the
CC modulation of NHP expression.
XX
SQ Sequence 945 BP; 323 A; 165 C; 202 G; 255 T; 0 other;

Query Match: 100.0%; Score 945; DB 22; Length 945;
Best Local Similarity 100.0%; Pred. NC. 4e-245;
Matches 945; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGAAAAGTATGAAAATATAGCTAAGACTGGAGAAGGGTCTTATGGGGTTGATTCAA 60
Db 1 ATGGAAAAGTATGAAAATATAGCTAAGACTGGAGAAGGGTCTTATGGGGTTGATTCAA 60
Qy 61 TGCAGAAACAAACCTCTGGACAAAGTAGTAGTGTGTTAAAAAATTTTGGAAATCTGAAGAT 120
Db 61 TGCAGAAACAAACCTCTGGACAAAGTAGTAGTGTGTTAAAAAATTTTGGAAATCTGAAGAT 120
Qy 121 GATCCTGTTGTTAAGAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTTAAACAT 180
Db 121 GATCCTGTTGTTAAGAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTTAAACAT 180
Qy 181 CCAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAGGAAAATGCAATTTAGTTT 240
Db 181 CCAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAGGAAAATGCAATTTAGTTT 240
Qy 241 GAATACCTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGAGATTGCTGAT 300
Db 241 GAATACCTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGAGATTGCTGAT 300
Qy 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCATATACAT 360
Db 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCATATACAT 360
Qy 361 AACTGTATTCACAGAGATATAAAACCTGAAAATATTTCTAATACTAAGCAAGGAATAATC 420
Db 361 AACTGTATTCACAGAGATATAAAACCTGAAAATATTTCTAATACTAAGCAAGGAATAATC 420
Qy 421 AAGATTGTGACTTCGGGTTTGCACAAATTTCTGATTCAGGAGATGCCTACACCGATTAT 480
Db 421 AAGATTGTGACTTCGGGTTTGCACAAATTTCTGATTCAGGAGATGCCTACACCGATTAT 480

Qy 481 GTAGCTACGAGATGGTACCGAGCTCCTGAACCTTCTTGTGGGAGATACCTAGTATGTTCT 540
Db 481 GTAGCTACGAGATGGTACCGAGCTCCTGAACCTTCTTGTGGGAGATACCTAGTATGTTCT 540
Qy 541 TCAGTCGATATATGGGCTATTGTTGTTGTTTTCAGAGCTCCTGACAGGCCAGCCACTG 600
Db 541 TCAGTCGATATATGGGCTATTGTTGTTGTTTTCAGAGCTCCTGACAGGCCAGCCACTG 600
Qy 601 TGGCTGGAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGAAAATTA 660
Db 601 TGGCTGGAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGAAAATTA 660
Qy 661 ATCCCAAGACATCAATCAATCTTTAAAAGTAAACGGGTTTTTCCATGGCATCAGTATACCT 720
Db 661 ATCCCAAGACATCAATCAATCTTTAAAAGTAAACGGGTTTTTCCATGGCATCAGTATACCT 720
Qy 721 GAGCCAGAAGACATGGAAACTCTTGAGGAAAAGTTCTCAGATGTTTCATCTGTGGCTCTG 780
Db 721 GAGCCAGAAGACATGGAAACTCTTGAGGAAAAGTTCTCAGATGTTTCATCTGTGGCTCTG 780
Qy 781 AACTTCATGAAGGGGTGTTGAAGATGAATCCAGATGACAGATTAACTGTTCCCAACTC 840
Db 781 AACTTCATGAAGGGGTGTTGAAGATGAATCCAGATGACAGATTAACTGTTCCCAACTC 840
Qy 841 CTGGAGAGCTCCTACTTTGATTTCTTTCAAGAGGCCCAAAATTTAAAGAAAAGCACGTAAT 900
Db 841 CTGGAGAGCTCCTACTTTGATTTCTTTCAAGAGGCCCAAAATTTAAAGAAAAGCACGTAAT 900
Qy 901 GAAGGAAGAAACAGAGAGACGCCCAACAGGTACTTTCGGCTCAAAAGT 945
Db 901 GAAGGAAGAAACAGAGAGACGCCCAACAGGTACTTTCGGCTCAAAAGT 945

RESULT 2
AAD03814
ID AAD03814 standard; cDNA; 972 BP.
XX
AC AAD03814;
XX
DT 19-JUN-2001 (first entry)
XX
DE Human kinase cDNA #3.
XX
KW Human; kinase; gene therapy; bioreactor; mental disorder;
KW biological disorder; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 1..972
FT /*tag= a
FT /product= "Human kinase #3"
FT /note= "The coding region does not include stop codon"
FT /partial
XX
PN WO200123579-A1.
XX
PD 05-APR-2001.
XX
PF 27-SEP-2000; 2000WO-US26621.
XX
PR 28-SEP-1999; 99US-0156511.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
XX
DR WPI; 2001-266166/27.
DR P-PSDB; AAE00495.
XX
PT New isolated human kinase polynucleotide useful for generating
PT antibodies, as reagents in diagnostic assays and for screening for

PT compounds useful for treating mental, biological or medical diseases -

XX Disclosure; Page 29-30; 38pp; English.

CC The present sequence is a cDNA encoding novel human protein (NHP) known as human kinase. The human kinases share structural similarity with animal kinases, more particularly serine or threonine protein kinases. Human kinase cDNA is useful for the detection of mutant human kinase for the diagnosis of disease, and also as a therapeutic. It is useful for screening drugs effective in the treatment of symptomatic or phenotypic manifestations perturbing the normal function of NHP in the body. The NHP nucleotide sequences are useful for generation of antibodies, as reagents in diagnostic assays, for the identification of other cellular gene products related to human kinases, and as reagents in assays for screening compounds that are useful for treating mental, biological or medical disorders.

CC NHP oligonucleotides are used as probes. The labelled NHP probes are useful for screening human genomic library for identifying polymorphisms and as primers in amplification assays to detect mutations within the exons, introns and splice sites that can be used in diagnostics and pharmacogenomics. Nucleotide construct encoding NHP products are used to genetically engineer cells in vivo that functions as bioreactors in the body delivering a continuous supply of NHP to the body. Nucleotide constructs encoding functional NHPs are used in gene therapy for the modulation of NHP expression.

XX

SQ Sequence 972 BP; 323 A; 172 C; 214 G; 263 T; 0 other;

Query Match 98.3%; Score 929; DB 22; Length 972;
Best Local Similarity 100.0%; Pred. No. 8.5e-241;
Matches 929; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAAGAAGTATGAAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGGTGTATTCAA 60
D5 |||||||
QY 1 ATGGAAGAAGTATGAAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGGTGTATTCAA 60
D5 |||||||

QY 61 TGCAGAAACAAACCTCTGGACAAGTAGTAGCTGTTAAAAAATTTGTGGAACTCTGAAGAT 120
D5 |||||||

QY 61 TGCAGAAACAAACCTCTGGACAAGTAGTAGCTGTTAAAAAATTTGTGGAACTCTGAAGAT 120
D5 |||||||

QY 121 GATCCTGTGTTAAGAAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAAT 180
D5 |||||||

QY 121 GATCCTGTGTTAAGAAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAAT 180
D5 |||||||

QY 181 CCAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAGGAAATGCATTTAGTTT 240
D5 |||||||

QY 181 CCAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAGGAAATGCATTTAGTTT 240
D5 |||||||

QY 241 GAATCTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGAGGTGCTGAT 300
D5 |||||||

QY 241 GAATCTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGAGGTGCTGAT 300
D5 |||||||

QY 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCTATACAT 360
D5 |||||||

QY 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCTATACAT 360
D5 |||||||

QY 361 AACTGTATTCACAGAGATATAAAACCTGAAATATTTCTAATACTAAGCAAGGAATATC 420
D5 |||||||

QY 361 AACTGTATTCACAGAGATATAAAACCTGAAATATTTCTAATACTAAGCAAGGAATATC 420
D5 |||||||

QY 421 AAGATTGTGACTTCGGGTTTGCACAAAATCTGATTCAGAGAGATGCCTACACCGATTAT 480
D5 |||||||

QY 421 AAGATTGTGACTTCGGGTTTGCACAAAATCTGATTCAGAGAGATGCCTACACCGATTAT 480
D5 |||||||

QY 481 GTAGCTACGAGATGGTACCGAGCTCCTGAACCTTCTTGTGGAGATACCTAGTATGGTCT 540
D5 |||||||

QY 481 GTAGCTACGAGATGGTACCGAGCTCCTGAACCTTCTTGTGGAGATACCTAGTATGGTCT 540
D5 |||||||

QY 541 TCAGTCGATATATGGGCTATTGGTTGTGTTTTTGCAGAGCTCCTGACAGGCCAGCCACTG 600
D5 |||||||

QY 541 TCAGTCGATATATGGGCTATTGGTTGTGTTTTTGCAGAGCTCCTGACAGGCCAGCCACTG 600
D5 |||||||

QY 601 TGGCCTGGAAAATCAGATGTGGACCACTTTATCTGATAATCAGAACACTAGGAAAATTA 660
D5 |||||||

D5 601 TGGCCTGGAAAATCAGATGTGGACCACTTTATCTGATAATCAGAACACTAGGAAAATTA 660
D5 |||||||

QY 661 ATCCCAAGACATCAATCAATCTTTAAAAGTAACGGGTTTTTCCATGGCATCAGTATACCT 720
D5 |||||||

D5 661 ATCCCAAGACATCAATCAATCTTTAAAAGTAACGGGTTTTTCCATGGCATCAGTATACCT 720
D5 |||||||

QY 721 GAGCCAGAGACATGGAACCTCTTGAGGAAAAGTTCTCAGATGTTTCATCCTGTGGCTCTG 780
D5 |||||||

D5 721 GAGCCAGAGACATGGAACCTCTTGAGGAAAAGTTCTCAGATGTTTCATCCTGTGGCTCTG 780
D5 |||||||

QY 781 AACTTCATGAAGGGGTGCTCTGAAGATGAATCCAGATGACAGATTAACTGTTCCCACTC 840
D5 |||||||

D5 781 AACTTCATGAAGGGGTGCTCTGAAGATGAATCCAGATGACAGATTAACTGTTCCCACTC 840
D5 |||||||

QY 841 CTGGAGAGCTCTACTTTTGATTCTTTTCAAGAGGCCCCAAATTAAGAAAGACGCTAAT 900
D5 |||||||

D5 841 CTGGAGAGCTCTACTTTTGATTCTTTTCAAGAGGCCCCAAATTAAGAAAGACGCTAAT 900
D5 |||||||

QY 901 GAAGGAAGAAACAGAAAGACGCCAACAGGT 929
D5 |||||||

D5 901 GAAGGAAGAAACAGAAAGACGCCAACAGGT 929
D5 |||||||

RESULT 3
AAD30557
ID AAD30557 standard; cDNA; 1790 BP.
XX
AC AAD30557;
XX 21-MAY-2002 (first entry)
XX
DE Human kinase polypeptide (PKIN-10) cDNA.
XX
KW Human; kinase polypeptide; PKIN-10; gene therapy; Addison's disease; leukaemia; immune disorder; lymphoma; melanoma; developmental disorder; acquired immunodeficiency syndrome; AIDS; allergy; anaemia; hypertension; asthma; Crohn's disease; rheumatoid arthritis; bursitis; atherosclerosis; cirrhosis; hepatitis; psoriasis; Cushing's syndrome; cytostatic; cancer; cholestasis; cardiac; cardiovascular disorder; Niemann-Pick's disease; lipid disorder; fatty liver; Gaucher's disease; myocardial infarction; drug screening; transgenic animal; antiinflammatory; hepatotropic; hypotensive; anti-HIV; enzyme; ss.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
FT CDS 239..1267
FT /*tag= a
FT /product= "Human PKIN-10"
XX
PN WO200208399-A2.
XX
PD 31-JAN-2002.
XX
PF 20-JUL-2001; 2001WO-US23092.
XX
PR 21-JUL-2000; 2000US-220038P.
PR 28-JUL-2000; 2000US-222112P.
PR 04-AUG-2000; 2000US-222831P.
PR 11-AUG-2000; 2000US-224729P.
XX
PA (INCY-) INCYTE GENOMICS INC.
PA (THOR/) THORNTON M.
XX
PI Thornton M, Yue H, Khan FA, Gururajan R, Hafalia AJA, Walia NK;
PI Patterson C, Ramkumar J, Gandhi AR, Policky JL, Baughn MR;
PI Tribouley CM, Bandman O, Nguyen DB, Lu Y, Burford N, Lal P;
PI Ding L, Yao MG, Elliott VS, Recipon SA, Kearney L, Lu DAM;
PI Greenwald SR, Tang YT, Xu Y, Walsh RT, Gietzen KJ, Yang J;
PI Hillman JL;

XX WPI; 2002-206083/26.
DR P-PSDB; AAE19152.
XX
PT New human kinase polypeptide, useful in diagnosis, prevention and
PT treatment of cancer, immune disorder, growth and developmental
PT disorder, cardiovascular disorder and lipid disorder -
XX
XX
PS Claim 5; Page 183; 196pp; English.
XX
CC The present invention relates to an isolated human kinase polypeptide
CC (PKIN) or a biologically active/immunogenic fragment of PKIN. PKIN is
CC useful for diagnosing, treating and preventing cancer (e.g., leukaemia,
CC lymphoma, melanoma), an immune disorder (e.g., acquired immunodeficiency
CC syndrome (AIDS), Addison's disease, allergy, anaemia, asthma, Crohn's
CC disease, rheumatoid arthritis), a growth and developmental disorder (e.g.
CC bursitis, cirrhosis, hepatitis, psoriasis, Cushing's syndrome), a
CC cardiovascular disorder (e.g., atherosclerosis, hypertension, myocardial
CC infarction), and a lipid disorder (e.g., fatty liver, cholestasis,
CC Gaucher's disease, Niemann-Pick's disease). PKIN is useful in a number of
CC drug screening techniques and to analyse the proteome of a tissue or cell
CC type. PKIN is useful for creating knockin humanised animals or transgenic
CC animals to model human diseases, in somatic or germline gene therapy, to
CC generate a transcript image of a tissue or cell type, for detecting
CC differences in the chromosomal location due to translocation, inversion,
CC etc., among normal, carrier or affected individuals, and as hybridisation
CC probes for mapping naturally occurring genomic sequences. PKIN is useful
CC in southern or northern analysis, dot blot or other membrane-based
CC technologies, in PCR technologies, in dipstick, pin, multiformat enzyme
CC linked immunosorbent (ELISA)-like assays and in microarrays utilising
CC fluids or tissues from patients to detect altered PKIN expression. The
CC present sequence is human PKIN-10 cDNA.
XX
SQ Sequence 1790 BP; 581 A; 328 C; 366 G; 515 T; 0 other;

Query Match 98.1%; Score 927; DB 24; Length 1790;
Best Local Similarity 100.0%; Pred. NO. 3.7e-240;
Matches 927; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAAAGTATGAAATATTAGCTAAGACTGGAGAAGGGTCTTATGGGTTGTATTCAA 60
DB 239 ATGGAAAGTATGAAATATTAGCTAAGACTGGAGAAGGGTCTTATGGGTTGTATTCAA 298
QY 61 TGCAGAAACAAACCTCTGGACAAGTAGTAGCTGTTAAATAATTTGTGGAATCTGAAGAT 120
DB 299 TGCAGAAACAAACCTCTGGACAAGTAGTAGCTGTTAAATAATTTGTGGAATCTGAAGAT 358
QY 121 GATCCTGTTGTTAAGAAATAGCACTAAGAGAAATACGTTATGTTGAAGCAATTAACAAT 180
DB 359 GATCCTGTTGTTAAGAAATAGCACTAAGAGAAATACGTTATGTTGAAGCAATTAACAAT 418
QY 181 CCAATCTTGTGACCTCATCGAGGTGTTCAGGAGAAAAGGAATGCAATTGATTTT 240
DB 419 CCAATCTTGTGACCTCATCGAGGTGTTCAGGAGAAAAGGAATGCAATTGATTTT 478
QY 241 GAATACCTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCCAATGGAGTTGCTGAT 300
DB 479 GAATACCTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCCAATGGAGTTGCTGAT 538
QY 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCATATACAT 360
DB 539 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCATATACAT 598
QY 361 AACTGTATTCACAGAGATATAAAACCTGAAATATTCTTAATAACTAAGCAAGGAATAATC 420
DB 599 AACTGTATTCACAGAGATATAAAACCTGAAATATTCTTAATAACTAAGCAAGGAATAATC 658
QY 421 AAGATTTGTGACITTCGGGTTTGCACAAATCTGATTCAGGAGATGCTACACCGATTAT 480
DB 659 AAGATTTGTGACTTCGGGTTTGCACAAATCTGATTCAGGAGATGCTACACCGATTAT 718
QY 481 GTAGCTACGAGATGGTACCGAGCTCCTGAACITCTTTGTGGGAGATACACTCAGTATGTTCT 540

DB 719 GTAGCTACGAGATGGTACCGAGCTCCTGAACCTTCTTGTGGAGATACTCAGTATGGTTCT 778
QY 541 TCAGTCGATATATGGGCTATGTTGTTGTTTGGAGAGCTCCTGACAGGGCCAGCCACTG 600
DB 779 TCAGTCGATATATGGGCTATGTTGTTTGGAGAGCTCCTGACAGGGCCAGCCACTG 838
QY 601 TGGCCTGGAAAAATCAGATGTGGACCAACTTTTATCTGATAATCAGAACACTAGGAAAAATTA 660
DB 839 TGGCCTGGAAAAATCAGATGTGGACCAACTTTTATCTGATAATCAGAACACTAGGAAAAATTA 898
QY 661 ATCCCAAGACATCAATCAATCTTTAAAAAGTAACGGGTTTTTCCATGGCATCAGTATACCT 720
DB 899 ATCCCAAGACATCAATCAATCTTTAAAAAGTAACGGGTTTTTCCATGGCATCAGTATACCT 958
QY 721 GAGCCAGAAGACATGGAAACTCTTGAGGAAAAAGTTCTCAGATGTTTCATCCGTGGCTCTG 780
DB 959 GAGCCAGAAGACATGGAAACTCTTGAGGAAAAAGTTCTCAGATGTTTCATCCGTGGCTCTG 1018
QY 781 AACTTCATGAAGGGGTGTTCTGAAGATGAATCCAGATGACAGATTAAACCTGTTCCCAACTC 840
DB 1019 AACTTCATGAAGGGGTGTTCTGAAGATGAATCCAGATGACAGATTAAACCTGTTCCCAACTC 1078
QY 841 CTGGAGAGCTCCTACTTTTGATTCTTTTCAAGAGGCCCAAAATTAAGAAAAAGCACGTAAT 900
DB 1079 CTGGAGAGCTCCTACTTTTGATTCTTTTCAAGAGGCCCAAAATTAAGAAAAAGCACGTAAT 1138
QY 901 GAAGGAAGAAACAGAGACGCGCCAAACAG 927
DB 1139 GAAGGAAGAAACAGAGACGCGCCAAACAG 1165

RESULT 4
ABX08936
CD ABX08936 standard; cDNA; 1678 BP.
XX AC ABX08936;
XX DT 21-JAN-2003 (first entry)
XX cDNA encoding human DITHP protein #6.
KW Human; ss; gene; diagnostic and therapeutic; DITHP; cancer;
KW arteriosclerosis atherosclerosis; psoriasis; asthma; autoimmune thrombocytopenia;
KW autoimmune; inflammatory; anaemia; diabetes mellitus; glomerulonephritis; gout; stroke;
KW Crohn's disease; rheumatoid arthritis; uveitis; AIDS; allergy;
KW multiple sclerosis; acquired immunodeficiency disease; neurological disorder; epilepsy;
KW Alzheimer's disease; dementia; mental retardation; gastrointestinal;
KW Parkinson's disease; ulcer; cirrhosis; reproductive; infertility;
KW endometriosis; endocrine disorder; hyperparathyroidism; hyperlipidemia;
KW hypercholesterolaemia; hypoglycaemia; obesity; Reiter's syndrome;
KW connective tissue disorder; osteoporosis; infection.
XX Homo sapiens.
XX WO200279473-A2.
PN 10-OCT-2002.
XX 09-JAN-2002; 2002WO-US01009.
PF 12-JAN-2001; 2001US-261622P.
XX 16-JAN-2001; 2001US-261864P.
PR 16-JAN-2001; 2001US-261865P.
PR 17-JAN-2001; 2001US-262164P.
PR 17-JAN-2001; 2001US-262207P.
PR 17-JAN-2001; 2001US-262208P.
PR 17-JAN-2001; 2001US-262209P.
PR 17-JAN-2001; 2001US-262215P.
PR 18-JAN-2001; 2001US-263102P.
PR 19-JAN-2001; 2001US-262599P.
PR 19-JAN-2001; 2001US-263063P.
PR 19-JAN-2001; 2001US-263064P.

PR 19-JAN-2001; 2001US-263065P.
PR 19-JAN-2001; 2001US-263069P.
PR 19-JAN-2001; 2001US-263077P.
PR 19-JAN-2001; 2001US-263329P.
PR 19-JAN-2001; 2001US-263330P.
XX (INCY-) INCYTE GENOMICS INC.
XX
PI Panzer SR, Lincoln SE, Altus CM, Dufour GE, Hillman JL, Jones AL;
PI Dam TC, Liu TF, Harris B, Flores V, Daffo A, Marwaha R, Chen AL;
PI Chang SC, Gerstin EH, Peralta CH, David MH, Lewis SA;
XX
DR WPI; 2003-040680/03.
DR P-PSDB; ABU05290.
XX
PT New human diagnostic and therapeutic (DITHP) polynucleotides and
PT polypeptides, useful for diagnosing, preventing or treating diseases,
PT e.g. cancer, AIDS, Parkinson's disease, or autoimmune/inflammatory
PT diseases -
XX
PS Claim 1; Page 267; 33lpp; English.
XX
CC This invention relates to the cDNA and protein sequences of fifty six
CC polynucleotides for diagnostics and therapeutics (DITHP). The human
CC DITHP polynucleotides and polypeptides are useful for diagnosing,
CC preventing or treating diseases associated with, as well as effects of
CC exogenous compounds, on the expression of human molecules, such as cell
CC proliferative diseases (e.g. cancer, arteriosclerosis atherosclerosis,
CC psoriasis, primary thrombocytopenia), autoimmune/inflammatory diseases
CC (e.g. anaemia, asthma, autoimmune thyroiditis, Crohn's disease, diabetes
CC mellitus, glomerulonephritis, gout, multiple sclerosis, rheumatoid
CC arthritis, uveitis, acquired immunodeficiency disease; AIDS; allergies;
CC neurological disorders (e.g. stroke, Alzheimer's disease, dementia,
CC mental retardation, Parkinson's disease, epilepsy), gastrointestinal
CC (e.g. ulcer, cirrhosis), reproductive (e.g. infertility, endometriosis),
CC endocrine disorders (e.g. hyperparathyroidism, hypercholesterolaemia,
CC hypoglycaemia, hyperlipidemia, obesity), connective tissue disorders
CC (e.g. osteoporosis, Reiter's syndrome), or infections (e.g. bacterial,
CC viral, fungal, parasitic, protozoal). The DITHP sequences may be used to
CC generate hybridisation probes useful in chromosomal mapping of naturally
CC occurring genomic sequences. They are also useful in designing probes
CC useful in diagnostic assays. The polynucleotides are useful as molecular
CC weight markers, or as antigen to elicit an immune response. The
CC present sequence represents a human diagnostics and therapeutics (DITHP)
CC cDNA sequence of the invention.
XX
SQ Sequence 1678 BP; 541 A; 308 C; 336 G; 493 T; 0 other;

Query Match 91.9%; Score 868.8; DB 25; Length 1678;
Best Local Similarity 99.2%; Pred. No. 1.9e-224;
Matches 925; Conservative 0; Mismatches 2; Indels 5; Gaps 5;

QY 1 ATGGAAAAGTATGAAAAATTACCTAAGACTGGAGAGGGTCTTATGGGGTGTATTCAA 60
DB 131 ATGGAGAAGTATGAAAAATTACCTAAGACTGGAGAGGGTCTTATGGGGTGTATTCAA 190
QY 61 TGCACA-AACAAAACCTCTGGACAAGTAGCTGTTAAAAAATTGTGGAACTGAAGA 119
DB 191 TGCAGATAACAAAACCTCTGGACAAGTAGCTGTTAAAAAATTGTGGAACTGAAGA 250
QY 120 TGATCCTGTTGTTAAGAAAATAGCAGTAAGAGAAATACGTA-TGTTGAAGCAATTAAAC 178
DB 251 TGATCCTGTTGTTAAGAAAATAGCAGTAAGAGAAATACGTA-TGTTGAAGCAATTAAAC 310
QY 179 ATCCAAATCTTGTGAACCTCATCGAGGTGTTTACGAGAAAAAGGAAAAATGCAATTTAGTT 238
DB 311 ATCCAAATCTTGTGAACCTCATCGAGGTGTTTACGAGAAAAAGGAAAAATGCAATTTAGTT 370
QY 239 TTGAATACTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGAGTTGCTG 298
DB 371 TTGAATACTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGAGTTGCTG 430
QY 299 ATGGAGTGATCAAAAAGCGTATTATGGCAACACTTCAAGCTCTTAAATTTCTGTCATATAC 358

Db 431 ATGGAGTGATCAAAAAGCGTATTATGCAAAACACTTCAAGCTCTTAATTTCTGTCAATAC 490
QY 359 ATAACCTGTATTCACAGAGATATAAAACCTGAAAAATATCTTAATAACTAAGCAAGGAATAA 418
DB 491 ATAACCTGTATTCACAGAGATATAAAACCTGAAAAATATCTTAATAACTAAGCAAGGAATAA 550
QY 419 TCAAGATTTGTGACTTCGGGTTTGCACAAAATCTGATTCAGGAGATGCCTACACCGATT 478
DB 551 TCAAGATTTGTGACTTCGGGTTTGCACAAAATCTGATTCAGGAGATGCCTACACCGATT 610
QY 479 ATGTAGCTA-CGAGATGTTACCGAGCT-CCTGAACCTTCTGTGGGAGATACT-CAGTATG 535
DB 611 ATGTAGCTAGGAGATGTTACCGAGCTCCCTGPAACCTTCTGTGGGAGATACTCCAGTATG 670
QY 536 GTTCTTCAGTCGATATATGGGCTATTGGTTGTGTTTTTTCAGAGAGCTCCTGACAGCCAGC 595
DB 671 GTTCTTCAGTCGATATATGGGCTATTGGTTGTGTTTTTTCAGAGAGCTCCTGACAGCCAGC 730
QY 596 CACTGTGGCCCTGGAAAAATCAGATGTGGACCAACTTTTATCTGATAATCAGAACACTAGGAA 655
DB 731 CACTGTGGCCCTGAAAAATCAGATGTGGACCAACTTTTATCTGATAATCAGAACACTAGGAA 790
QY 656 AATTATATCCCAAGACATCAATCAATCTTTAAAAAGTAACGGGTTTTTCCATGGCATCAGTA 715
DB 791 AATTATATCCCAAGACATCAATCAATCTTTAAAAAGTAACGGGTTTTTCCATGGCATCAGTA 850
QY 716 TACCTGAGCCAGAGACATGGAACCTCTTGAGGAAAAAGTTCTCAGATGTTTCATCTCTGTTGG 775
DB 851 TACCTGAGCCAGAGACATGGAACCTCTTGAGGAAAAAGTTCTCAGATGTTTCATCTCTGTTGG 910
QY 776 CTCTGAACCTTCATGAAGGGGTGTTGAAGATGAATCCAGATGACAGATTAACCTGTTCCTC 835
DB 911 CTCTGAACCTTCATGAAGGGGTGTTGAAGATGAATCCAGATGACAGATTAACCTGTTCCTC 970
QY 836 AACTCCTGGAGAGCTCCTACTTTGATTTCTTTTCAAGAGGGCCCAAAATTAAGAAGAAAGCAC 895
DB 971 AACTCCTGGAGAGCTCCTACTTTGATTTCTTTTCAAGAGGGCCCAAAATTAAGAAGAAAGCAC 1030
QY 896 GTAATGAAGGAAGAAACAGAGACGCGCCAAACAG 927
DB 1031 GTAATGAAGGAAGAAACAGAGACGCGCCAAACAG 1062

RESULT 5
AAS06725
ID AAS06725 standard; cDNA; 1083 BP.
XX
AC AAS06725;
XX
DT 12-SEP-2001 (first entry)
XX
DE Polynucleotide sequence encoding human protein kinase #25.
XX
KW Human; protein kinase; PK; STK; cancer; cardiovascular disease;
KW metabolic disorder; immune related disease; neurological disorder;
KW neurodegenerative disorder; inflammatory disorder; infectious disease;
KW reproductive disorder; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200138503-A2.
XX
PD 31-MAY-2001.
XX
PF 22-NOV-2000; 2000WO-US32085.
XX
PR 24-NOV-1999; 99US-0167482.
XX
PA (SUGEN-) SUGEN INC.
PI Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
PI Flanagan P, Clary D;

XX WPI; 2001-343950/36.
DR P-PSDB; AAU03525.
XX
PT Nucleic acids encoding human kinase polypeptides, useful for preventing
PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and
PT neuronal-associated diseases, and microbial infections -
XX
PS Example 1; Figure 1; 433pp; English.
XX
CC AAS06701-AAS06757 encode for novel human protein kinases #1-57. The
CC novel protein kinases have been identified as members of the tyrosine
CC or serine/threonine kinase (PTK and STK) families. The polynucleotides
CC encoding protein kinases and the polypeptides may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate kinase expression. For example, they may be used to treat
CC cancers (especially cancers of haematopoietic origin), cardiovascular
CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
CC immune related diseases (e.g. rheumatoid arthritis), neurological
CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
CC disease (e.g. HIV), and reproductive disorders (e.g. infertility).
CC Additionally, polynucleotides encoding protein kinases may be
CC used for gene therapy and as DNA probes in diagnostic assays.
CC The protein kinase polypeptides may be used as antigens in the production
CC of antibodies against the protein kinases and in assays to identify
CC modulators of protein kinase expression and activity.
XX
SQ Sequence 1083 BP; 366 A; 199 C; 225 G; 293 T; 3 other;

Query Match 91.3%; Score 863; DB 22; Length 1083;
Best Local Similarity 94.5%; Pred. No. 5.9e-223;
Matches 927; Conservative 0; Mismatches 0; Indels 54; Gaps 1;

Qy 2 ATGGAAAAGTATGAAAATTAGCTAAGACTGGAGAAGGGTCTTATGGGGTTGTATTCAA 60
Db 1 ATGGAAAAGTATGAAAATTAGCTAAGACTGGAGAAGGGTCTTATGGGGTTGTATTCAA 60

Qy 61 TGCAGAAACAAACCTCTGGACAAGTAGTAGCTGTAAATAATTTGTGGAATCTGAAGAT 120
Db 61 TGCAGAAACAAACCTCTGGACAAGTAGTAGCTGTAAATAATTTGTGGAATCTGAAGAT 120

Qy 121 GATCCTGTTGTTAAGAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAT 180
Db 121 GATCCTGTTGTTAAGAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAT 180

Qy 181 CCAATCTTGTGAACCTCACTCGAGGTGTTTCAAGAGAAAAGCAATTCGATTTT 240
Db 181 CCAATCTTGTGAACCTCACTCGAGGTGTTTCAAGAGAAAAGCAATTCGATTTT 240

Qy 241 GAATACGTGTATACATACATCTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT 300
Db 241 GAATACGTGTATACATACATCTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT 300

Qy 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTATATACAT 360
Db 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTATATACAT 360

Qy 361 AACTGTATTACAGAGATATAAAACCTGAAATATTTCTAATACTAAGCAAGGAATATC 420
Db 361 AACTGTATTACAGAGATATAAAACCTGAAATATTTCTAATACTAAGCAAGGAATATC 420

Qy 421 AAGATTGTGACTTCGGGTTTGACAAATTCGA----- 454
Db 421 AAGATTGTGACTTCGGGTTTGACAAATTCGAATTCACATGTTGGGAGAAATTGAC 480

Qy 455 -----TTCCAGAGATGCCTACACCGATTATGTAGCT 486
Db 481 TTACCTGATCTGATTGATGCAATTTGCAGTTCCAGAGATGCCTACACCGATTATGTAGCT 540

Qy 487 ACGAGATGGTACCGAGCTCCTGAACCTTCTTGTGGAGATACCTCAGTATGTTCTTCAGTC 546
Db 541 ACGAGATGGTACCGAGCTCCTGAACCTTCTTGTGGAGATACCTCAGTATGTTCTTCAGTC 600

QY 547 GATATATGGGCTATTGGTGTGTTTTCAGAGCTCTGACAGGCCAGCCACTGTGGCCT 606
Db 601 GATATATGGGCTATTGGTGTGTTTTCAGAGCTCTGACAGGCCAGCCACTGTGGCCT 660

QY 607 GGAAATCAGATGTGGACCAACTTTATCTGATATATCAGAACACTAGGAAATTAATCCCA 666
Db 661 GGAAATCAGATGTGGACCAACTTTATCTGATATATCAGAACACTAGGAAATTAATCCCA 720

QY 667 AGACATCAATCAATCTTTAAAGTAACGGGTTTTCATGGCATCAGTATACCTGAGCCA 726
Db 721 AGACATCAATCAATCTTTAAAGTAACGGGTTTTCATGGCATCAGTATACCTGAGCCA 780

QY 727 GAAGACATGGAAACTCTTGAGGAAAAGTTCTCAGATGTTTCATCCTGTGGCTCTGAATTC 786
Db 781 GAAGACATGGAAACTCTTGAGGAAAAGTTCTCAGATGTTTCATCCTGTGGCTCTGAATTC 840

QY 787 ATGAAGGGGTGTTCTGAAGATGAATCCAGATGACAGATTAACCTGTTCCCAACTCCTGGAG 846
Db 841 ATGAAGGGGTGTTCTGAAGATGAATCCAGATGACAGATTAACCTGTTCCCAACTCCTGGAG 900

QY 847 AGCTCCTACTTTGATTCTTTCAAGAGGCCCAATTAAGAAAAGCAACGTAATGAAGGA 906
Db 901 AGCTCCTACTTTGATTCTTTCAAGAGGCCCAATTAAGAAAAGCAACGTAATGAAGGA 960

QY 907 AGAAACAGAAAGACGCCAACAG 927
Db 961 AGAAACAGAAAGACGCCAACAG 981

RESULT 6
AAD03816
ID AAD03816 standard; cDNA; 1041 BP.
XX
AC AAD03816;
XX
DT 19-JUN-2001 (first entry)
XX
DE Human kinase cDNA #5.
XX
KW Human; kinase; gene therapy; bioreactor; mental disorder;
KW biological disorder; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1041
FT /tag= a
FT /product= "Human kinase #5"
FT /note= "The coding region does not include stop codon"
FT /partial
XX
PN WO200123579-A1.
XX
PD 05-APR-2001.
XX
PF 27-SEP-2000; 2000WO-US26621.
XX
PR 28-SEP-1999; 99US-0156511.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
XX
DR WPI; 2001-266166/27.
DR P-PSDB; AAE00494.
XX
PT New isolated human kinase polynucleotide useful for generating
PT antibodies, as reagents in diagnostic assays and for screening for
PT compounds useful for treating mental, biological or medical diseases -
XX
PS Claim 1; Page 31-32; 38pp; English.
XX

CC The present sequence is a cDNA encoding novel human protein
CC (NHP) known as human kinase. The human kinases share structural
CC similarity with animal kinases, more particularly serine or
CC threonine protein kinases. Human kinase cDNA is useful for the
CC detection of mutant human kinase for the diagnosis of disease,
CC and also as a therapeutic. It is useful for screening drugs
CC effective in the treatment of symptomatic or phenotypic
CC manifestations perturbing the normal function of NHP in the
CC body. The NHP nucleotide sequences are useful for generation of
CC antibodies, as reagents in diagnostic assays, for the
CC identification of other cellular gene products related to human
CC kinases, and as reagents in assays for screening compounds that
CC are useful for treating mental, biological or medical disorders.
CC NHP oligonucleotides are used as probes. The labelled NHP probes
CC are useful for screening human genomic library for identifying
CC polymorphisms and as primers in amplification assays to detect
CC mutations within the exons, introns and splice sites that can
CC be used in diagnostics and pharmacogenomics. Nucleotide construct
CC encoding NHP products are used to genetically engineer cells
CC in vivo that functions as bioreactors in the body delivering a
CC continuous supply of NHP to the body. Nucleotide constructs
CC encoding functional NHPs are used in gene therapy for the
CC modulation of NHP expression.

XX SQ Sequence 1041 BP; 344 A; 190 C; 228 G; 279 T; 0 other;

Query Match 89.8%; Score 839; DB 22; Length 1041;
Best Local Similarity 90.8%; Pred. No. 1.8e-216;
Matches 945; Conservative 0; Mismatches 0; Indels 96; Gaps 1;

QY 1 ATGGAAGAAGTATGAAAAATAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAA 60
DB 1 ATGGAAGAAGTATGAAAAATAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAA 60
QY 61 TGCAGAAACAAACCTCTGGCAAGTAGTAGCTGTAAATAATTTGTGGAATCTGAAGAT 120
DB 61 TGCAGAAACAAACCTCTGGCAAGTAGTAGCTGTAAATAATTTGTGGAATCTGAAGAT 120
QY 121 GATCCTGTGTTAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAT 180
DB 121 GATCCTGTGTTAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAT 180
QY 181 CCAATCTTGTGAACCTCATCGAGGTGTTCCAGGAGAAAGGAAATGCATTTAGTTTT 240
DB 181 CCAATCTTGTGAACCTCATCGAGGTGTTCCAGGAGAAAGGAAATGCATTTAGTTTT 240
QY 241 GAATACCTGTGATCATACACTTTTAATGAGCTGGAAGAAACCCAAATGGAGTTGCTGAT 300
DB 241 GAATACCTGTGATCATACACTTTTAATGAGCTGGAAGAAACCCAAATGGAGTTGCTGAT 300
QY 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGCATATACAT 360
DB 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGCATATACAT 360
QY 361 AACTGTATTCACAGAGATATAAAACCTGAAATATTTCTAATACTAAGCAAGGAATATC 420
DB 361 AACTGTATTCACAGAGATATAAAACCTGAAATATTTCTAATACTAAGCAAGGAATATC 420
QY 421 AAGATTGTGACTTCGGGTTTGCACAAATTTCTGATCCAGGAGATGCCACCGATTAT 480
DB 421 AAGATTGTGACTTCGGGTTTGCACAAATTTCTGATCCAGGAGATGCCACCGATTAT 480
QY 481 GTAGCTACGAGATGGTACCGAGCTCCTGAACCTTTGTTGGGAGATACCTAGTATGTTCT 540
DB 481 GTAGCTACGAGATGGTACCGAGCTCCTGAACCTTTGTTGGGAGATACCTAGTATGTTCT 540
QY 541 TCAGTCGATATATGGGCTATTGTTGTTTTCAGAGCTCCTGACAGGCGAGCCACTG 600
DB 541 TCAGTCGATATATGGGCTATTGTTGTTTTCAGAGCTCCTGACAGGCGAGCCACTG 600
QY 601 TGGCCTGGAAATCAGATGTGGACCAACTTTATCTGATTAATCAGAACACT----- 650
DB 601 TGGCCTGGAAATCAGATGTGGACCAACTTTATCTGATTAATCAGAACACTAGTAGAGACG 650

QY 651 ----- 650
DB 661 GGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTCTTGACGTCAAGTGATCCACCTGCC 720
QY 651 -----AGGAAATTAATCCCAAGACATCAATCAATCTTT 684
DB 721 GTAGCCTCTCAAAAGTCTGTAATTACAGGAAATTAATCCCAAGACATCAATCAATCTTT 780
QY 685 AAAAGTAACGGGTTTTCATGGCATCAGTATACCTGAGCAGAGACATGGAAATCTTT 744
DB 781 AAAAGTAACGGGTTTTCATGGCATCAGTATACCTGAGCAGAGACATGGAAATCTTT 840
QY 745 GAGGAAAAGTCTCAGATGTTTCATCCTGTGGCTCTGAACTTCATGAAGGGGTGCTGAAG 804
DB 841 GAGGAAAAGTCTCAGATGTTTCATCCTGTGGCTCTGAACTTCATGAAGGGGTGCTGAAG 900
QY 805 ATGAATCCAGATGACAGATTAACTGTTCCCAACTCCTGGAGAGCTCCTACTTTGATTCT 864
DB 901 ATGAATCCAGATGACAGATTAACTGTTCCCAACTCCTGGAGAGCTCCTACTTTGATTCT 960
QY 865 TTTCAAGAGGCCCAATTAAGAAAGAACAGCAGTAATGAAGAAAGAACAGAGACGCCAA 924
DB 961 TTTCAAGAGGCCCAATTAAGAAAGAACAGCAGTAATGAAGAAAGAACAGAGACGCCAA 1020
QY 925 CAGGTACTTCCGCTCAAAAGT 945
DB 1021 CAGGTACTTCCGCTCAAAAGT 1041

RESULT 7
AAD03813
ID AAD03813 standard; cDNA; 1068 BP.
XX
AC AAD03813;
DT 19-JUN-2001 (first entry)
XX
DE Human kinase cDNA #2.
XX
KW Human; kinase; gene therapy; bioreactor; mental disorder;
biological disorder; ss.
XX
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 1..1068
FT /*tag= a
FT /product= "Human kinase #2"
FT /note= "The coding region does not include stop codon"
FT /partial
PN WO200123579-A1.
XX
PD 05-APR-2001.
XX
PF 27-SEP-2000; 2000WO-US26621.
XX
PR 28-SEP-1999; 99US-0156511.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
XX
DR WPI; 2001-266166/27.
DR P-PSDS; AAE00491.
XX
PT New isolated human kinase polynucleotide useful for generating
PT antibodies, as reagents in diagnostic assays and for screening for
PT compounds useful for treating mental, biological or medical diseases -
XX
PS Disclosure; Page 28; 38pp; English.
XX

CC The present sequence is a cDNA encoding novel human protein
CC (NHP) known as human kinase. The human kinases share structural
CC similarity with animal kinases, more particularly serine or
CC threonine protein kinases. Human kinase cDNA is useful for the
CC detection of mutant human kinase for the diagnosis of disease,
CC and also as a therapeutic. It is useful for screening drugs
CC effective in the treatment of symptomatic or phenotypic
CC manifestations perturbing the normal function of NHP in the
CC body. The NHP nucleotide sequences are useful for generation of
CC antibodies, as reagents in diagnostic assays, for the
CC identification of other cellular gene products related to human
CC kinases, and as reagents in assays for screening compounds that
CC are useful for treating mental, biological or medical disorders.
CC NHP oligonucleotides are used as probes. The labelled NHP probes
CC are useful for screening human genomic library for identifying
CC polymorphisms and as primers in amplification assays to detect
CC mutations within the exons, introns and splice sites that can
CC be used in diagnostics and pharmacogenomics. Nucleotide construct
CC encoding NHP products are used to genetically engineer cells
CC in vivo that functions as bioreactors in the body delivering a
CC continuous supply of NHP to the body. Nucleotide constructs
CC encoding functional NHPs are used in gene therapy for the
CC modulation of NHP expression.

XX
SQ Sequence 1068 BP; 344 A; 197 C; 240 G; 287 T; 0 other;

Query Match 87.1%; Score 823; DB 22; Length 1068;
Best Local Similarity 90.6%; Pred. No. 3.8e-212;
Matches 929; Conservative 0; Mismatches 0; Indels 96; Gaps 1;

QY 2 ATGGAAGATGATGAAATATAGCTAAGACTGGAGAGGGCTCTTATGGGTTGTATTCAA 60
Db 1 ATGGAAGATGATGAAATATAGCTAAGACTGGAGAGGGCTCTTATGGGTTGTATTCAA 60
QY 61 TGCAGAAACAAACCTCTGACCAAGTAGTAGCTGTTAAATAATTTGTGGAATCTGAAGAT 120
Db 61 TGCAGAAACAAACCTCTGACCAAGTAGTAGCTGTTAAATAATTTGTGGAATCTGAAGAT 120
QY 121 GATCCTGTTGTTAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAT 180
Db 121 GATCCTGTTGTTAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAT 180
QY 181 CCATCTGTTGAACCTCATCGAGGTGTTTCCAGGAGAAAGGAAATGCAATTTAGTTT 240
Db 181 CCATCTGTTGAACCTCATCGAGGTGTTTCCAGGAGAAAGGAAATGCAATTTAGTTT 240
QY 241 GAATCTGTTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGGAGTTGCTGAT 300
Db 241 GAATCTGTTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGGAGTTGCTGAT 300
QY 301 GGAGTGATCAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGTCATATACAT 360
Db 301 GGAGTGATCAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGTCATATACAT 360
QY 361 AACTGTATTACAGAGATATAAAACCTGAAATATTTCTAATACTAAGCAAGCAATAATC 420
Db 361 AACTGTATTACAGAGATATAAAACCTGAAATATTTCTAATACTAAGCAAGCAATAATC 420
QY 421 AAGATTTGTGACTTCGGGTTTGCACAAATTTCTGATTTCCAGGAGATGCTTACACCGATTAT 480
Db 421 AAGATTTGTGACTTCGGGTTTGCACAAATTTCTGATTTCCAGGAGATGCTTACACCGATTAT 480
QY 481 GTAGCTACGAGATGGTACCGAGCTCTGAACTTCTTGTGGGAGATACCTCAGTATGTTCT 540
Db 481 GTAGCTACGAGATGGTACCGAGCTCTGAACTTCTTGTGGGAGATACCTCAGTATGTTCT 540
QY 541 TCAGTCGATATATGGGCTATTGGTTGTGTTTTTGCAGAGCTCCTGACAGGCCACCACTG 600
Db 541 TCAGTCGATATATGGGCTATTGGTTGTGTTTTTGCAGAGCTCCTGACAGGCCACCACTG 600
QY 601 TGGCCTGGAAATCAGATGTGGACCACTTTATCTGATATCAGAACACT----- 650
Db 601 TGGCCTGGAAATCAGATGTGGACCACTTTATCTGATATCAGAACACTAGTAGAGACG 660

QY 651 ----- 650
Db 661 GGGTTTCGCCATGTTGACCCAGGCTGGTCTCGAACTTTGACGTCAAGTGATCCACCTGCC 720
QY 651 -----AGGAAATTAATCCCAAGACATCAATCAATCTTT 684
Db 721 GTAGCCTCTCAAAGTCTGGATTAACAGGAAATTAATCCCAAGACATCAATCAATCTTT 780
QY 685 AAAAGTAACGGGTTTTTCCATGGCATCAGTATACCTGAGCCAGAGACATGGAACCTCTT 744
Db 781 AAAAGTAACGGGTTTTTCCATGGCATCAGTATACCTGAGCCAGAGACATGGAACCTCTT 840
QY 745 GAGGAAAGTTCTCAGATGTTTCATCTCTGTGGCTCTGAACTTCATGAAGGGGTCTCTGAAG 804
Db 841 GAGGAAAGTTCTCAGATGTTTCATCTCTGTGGCTCTGAACTTCATGAAGGGGTCTCTGAAG 900
QY 805 ATGAATCCAGATGACAGATTAACTGTTCCCAACTCTCTGGAGAGCTCCTACTTTGATTCT 864
Db 901 ATGAATCCAGATGACAGATTAACTGTTCCCAACTCTCTGGAGAGCTCCTACTTTGATTCT 960
QY 865 TTTCAAGAGGCCCAATTAATAAGAAAGACACGTAATGAAGGAAGAAACAGACAGCCCAA 924
Db 961 TTTCAAGAGGCCCAATTAATAAGAAAGACACGTAATGAAGGAAGAAACAGACAGCCCAA 1020
QY 925 CAGGT 929
Db 1021 CAGGT 1025

RESULT 8
AA164248
ID AA164248 standard; cDNA; 882 BP.
XX

AC AA164248;
XX
DT 08-MAR-2002 (first entry)
XX
DE Human Kinase 14257 cDNA.

XX Protein kinase; enzyme; cytostatic; osteopathic; hepatotropic;
KW antidiabetic; neuroprotective; antiarthritic; dermatological;
KW immunosuppressive; antiinflammatory; antithyroid; antipsoriatic;
KW ophthalmological; antiallergic; antiasthmatic; antiatherosclerotic;
KW hypotensive; vasotropic; antiarrhythmic; virucide; anorectic;
KW metabolic; immunomodulator; analgesic; cellular proliferative disorder;
KW cancer; acute lymphoblastic leukaemia; Hodgkin's disease;
KW bone metabolism disorder; osteoporosis; immune system disorder;
KW inflammatory; diabetes mellitus; osteoarthritis; asthma;
KW cardiovascular disorder; hypertension; coronary artery disease;
KW endothe-lal cell disorder; psoriasis; ss.

OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 5'UTR 1..128
FT /*tag= a
FT CDS 129..815
FT /*tag= b
FT /*product= "Protein kinase 14275"
FT 3'UTR 816..882
FT /*tag= c

XX WO200179488-A2.
PN
XX
PD 25-OCT-2001.
XX
PF 13-APR-2001; 2001WO-US12188.
XX
PR 13-APR-2000; 2000US-196910P.
XX (MILL-) MILLENNIUM PHARM INC.
XX

PI Kapeller-libermann R;
XX WPI; 2002-034355/04.
DR P-PSDB; AAG78547.
XX
PT New 14257 polypeptides (protein kinases), useful as diagnostic targets
PT and therapeutic agents for controlling cellular proliferative and/or
PT differentiative disorder, bone disorders, immune disorders and
PT cardiovascular disorders
XX
PS Claim 2; Fig 1A; 98pp; English.
XX
CC The invention relates to an isolated 14257 polypeptide and nucleic
CC acid encoding it. The 14257 protein is a protein kinase that acts as a
CC modulating agent in regulating a variety of cellular processes,
CC including cell proliferation, differentiation, growth and division.
CC The activity of the protein of the invention may be described as:
CC cytosolic; osteopathic; hepatotropic; antidiabetic; neuroprotective;
CC antiarthritic; dermatological; immunosuppressive; antiinflammatory;
CC antithyroid; antipsoriatic; ophthalmological; antiallergic;
CC antiasthmatic; antiatherosclerotic; hypotensive; vasotropic;
CC antiarrhythmic; virucide; anorectic; metabolic; immunomodulator and
CC analgesic. The protein of the invention may act as a novel diagnostic
CC target or therapeutic agent controlling certain disorders, for example
CC kinase-associated or other 14257-associated disorders. These may include
CC cellular proliferative disorders such as cancers e.g. acute lymphoblastic
CC leukemia or Hodgkin's disease. Other disorders include bone metabolism
CC disorders such as osteoporosis, disorders of the immune system, e.g.
CC inflammatory, diabetes mellitus, osteoarthritis and asthma. Proteins of
CC the invention may also be of use as therapeutic agents in cardiovascular
CC disorders such as hypertension and coronary artery disease, and some
CC endothelial cell disorders, including psoriasis. The current
CC sequence represents a human kinase 14257 cDNA.
XX
SQ Sequence 882 BP; 288 A; 150 C; 198 G; 245 T; 1 other;

Query Match 68.3%; Score 645.4; DB 24; Length 882;
Best Local Similarity 99.1%; Pred. No. 3.3e-164;
Matches 649; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 ATGGAAAGATGATAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGGTTGTATTCAA 60
DB 129 ATGGGAAAGATGATAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGGTTGTATTCAA 188
QY 61 TGCAGAAACAAACCTCTGGACAGTAGTAGCTGTAAATAAATTTGTGGATCTGAAGAT 120
DB 189 TGCAGAAACAAACCTCTGGACAGTAGTAGCTGTAAATAAATTTGTGGATCTGAAGAT 248
QY 121 GATCCTGTGTTAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAAACAT 180
DB 249 GATCCTATGTTAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAAACAT 308
QY 181 CCAATCTTGTGAACCTCATCGAGGTGTTCAAGAGAAAGGAAATGCAATTAGTTTTT 240
DB 309 CCAATCTTGTGAACCTCATCGAGGTGTTCAAGAGAAAGGAAATGCAATTAGTTTTT 368
QY 241 GAATCTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT 300
DB 369 GAATCTGTGATCATGCACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT 428
QY 301 GGAGTGATCAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGTATATACAT 360
DB 429 GGAGTGATCAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGTATATACAT 488
QY 361 AACTGTATTACAGAGATATAAAACCTGAAATATTCTAATAACTAAGCAAGGAATAATC 420
DB 489 AACTGTATTACAGAGATATAAAACCTGAAATATTCTAATAACTAAGCAAGGAATAATC 548
QY 421 AAGATTGTGACTTCGGGTTTGACAAATCTGATTCAGGAGATGCTACACCGATTAT 480
DB 549 AAGATTGTGACTTCGGGTTTGACAAATCTGATTCAGGAGATGCTACACCGATTAT 608
QY 481 GTAGCTACGAGATGGTACCGAGCTCTGAACTTCTTGTGGAGATACACTAGTGTCTT 540

DB 609 GTAGCTACGAGATGGTACCGAGCTCTGAAACTTCTTGTGGAGATACTCAGTATGTTCT 668
QY 541 TCAGTCGATATATGGGCTATTGTTGTGTTTTTGCAGAGCTCTGACAGGCCAGCCACTG 600
DB 669 TCAGTCGATATATGGGCTATTGTTGTGTTTTTGCAGAGCTCTGACAGGCCAGCCACTG 728
QY 601 TGGCCTGGAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACTAGGAA 655
DB 729 TGGCCTGGAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACTAGGTA 783

RESULT 9
AAF44669
ID AAF44669 standard; cDNA; 911 BP.
XX AAF44669;
XX 27-MAR-2001 (first entry)
DE Novei protein kinase cDNA, SEQ ID NO: 49.
XX Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
XX immunosuppressive; cardiant; renal; antiinflammatory; antiasthmatic;
XX dermatological; antidiabetic; antifertility; gene therapy; vaccine;
XX immune disorder; cardiovascular disease; neurodegenerative disease;
XX cancer; autoimmune disorder; stroke; inflammatory bowel disease;
XX inflammatory pelvic disease; multiple sclerosis; psoriasis; ss.
OS Homo sapiens.
XX WO200073469-A2.
PN 07-DEC-2000.
XX 26-MAY-2000; 2000WO-US14842.
PR 28-MAY-1999; 99US-0136503.
XX (SUGEN) SUGEN INC.
PA Plowman GD, Martinez R, Whyte D, Sudersanam S;
P- WPI; 2001-032161/04.
XX P-PSDB; AAB65642.
DR Nucleic acids encoding kinase polypeptides, useful for diagnosing and
XX treating immune-related diseases and disorders, cardiovascular disease,
XX neurodegenerative diseases and/or cancers -
PS Example 4; Fig 2; 310pp; English.
XX
CC The present sequence encodes a novel protein kinase. The nucleic acids
CC and the protein kinases they encode may be used in the treatment and
CC diagnosis of diseases associated with inappropriate kinase expression
CC such as immune-related diseases and disorders, cardiovascular disease,
CC neurodegenerative diseases and/or cancers. The nucleic acids and
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays. The kinase polypeptides may be used as antigens in the production
CC of antibodies of kinase expression and activity. Anti-kinase antibodies
CC and kinase antagonists may also be used to down regulate kinase
CC expression and activity. Diseases related to kinase expression and
CC activity include rheumatoid arthritis, atherosclerosis, autoimmune
CC disorders, complications of organ transplantation, myocardial infarction,
CC immune disorders, cardiomyopathies, strokes, renal failure,
CC oxidative stress related disorders, chronic inflammatory bowel disease,
CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,
CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and
XX reproductive disorders.
SQ Sequence 911 BP; 293 A; 182 C; 178 G; 258 T; 0 other;

Query Match 68.0%; Score 643; DB 22; Length 911;

Best Local Similarity 100.0%; Pred. No. 1.5e-163;
Matches 643; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 285 AAATGGAGTTCCTGATGGAGTGATCAAAAGCGTATTATGGCAAACTTCAAGCTCTTAA 344
DB 1 AAATGGAGTTCCTGATGGAGTGATCAAAAGCGTATTATGGCAAACTTCAAGCTCTTAA 60

QY 345 TTTCTGTCAATATACATAAATCTGATTTACAGAGATATAAAACCTGAAATATTCTAATAAC 404
DB 61 TTTCTGTCAATATACATAAATCTGATTTACAGAGATATAAAACCTGAAATATTCTAATAAC 120

QY 405 TAAGCAAGGAATATCAAGATTGTGACTTCGGGTTTGACAAATCTGATTCAGGAGA 464
DB 121 TAAGCAAGGAATATCAAGATTGTGACTTCGGGTTTGACAAATCTGATTCAGGAGA 180

QY 465 TGCCTACACCGATTATGTAGCTACGAGATGGTACCGAGCTCCTGAACTTCTTGTGGAGA 524
DB 181 TGCCTACACCGATTATGTAGCTACGAGATGGTACCGAGCTCCTGAACTTCTTGTGGAGA 240

QY 525 TACTCAGTATGGTTCCTCAGTCGATATATGGGCTATTGGTGTGTTTTCAGAGCTCCT 584
DB 241 TACTCAGTATGGTTCCTCAGTCGATATATGGGCTATTGGTGTGTTTTCAGAGCTCCT 300

QY 585 GACAGGCCAGCCACTGTGGCTGGAAATCAGATGTGGACCACTTATCTGATAATCAG 644
DB 301 GACAGGCCAGCCACTGTGGCTGGAAATCAGATGTGGACCACTTATCTGATAATCAG 360

QY 645 AACACTAGGAAAATTAATCCCAAGACATCAATCAATCTTTAAAGTAACGGGTTTTCCTCA 704
DB 361 AACACTAGGAAAATTAATCCCAAGACATCAATCAATCTTTAAAGTAACGGGTTTTCCTCA 420

QY 705 TGGCATCAGTATACCTGAGCCAGAGACATGGAACTCTTGAGGAAAAGTCTCAGATGT 764
DB 421 TGGCATCAGTATACCTGAGCCAGAGACATGGAACTCTTGAGGAAAAGTCTCAGATGT 480

QY 765 TCATCCTGTGGCTGTGAACCTCATGAAGGGGTGTCTGAAGATGAATCCAGATGACAGATT 824
DB 481 TCATCCTGTGGCTGTGAACCTCATGAAGGGGTGTCTGAAGATGAATCCAGATGACAGATT 540

QY 825 AACCTGTTCCCACTCCTGGAGAGCTCCTACTTGTGATCTTTTCAAGAGGCCCAATTA 884
DB 541 AACCTGTTCCCACTCCTGGAGAGCTCCTACTTGTGATCTTTTCAAGAGGCCCAATTA 600

QY 885 AAGAAAAGCACGTAATGAAGGAAGAAACAGAGACGCCAACAG 927
DB 601 AAGAAAAGCACGTAATGAAGGAAGAAACAGAGACGCCAACAG 643

RESULT 10
AAF44670
ID AAF44670 standard; cDNA; 2615 BP.

XX
AC AAF44670;
XX
DT 27-MAR-2001 (first entry)

XX
DE Novel protein kinase cDNA, SEQ ID NO: 50.

XX
KW Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
KW immunosuppressive; cardiant; renal; antiinflammatory; antiasthmatic;
KW dermatological; antidiabetic; antiinfertility; gene therapy; vaccine;
KW immune disorder; cardiovascular disease; neurodegenerative disease;
KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;
KW inflammatory pelvic disease; multiple sclerosis; psoriasis; ss.

OS Mus musculus.
XX
PN WO200073469-A2.
XX
PD 07-DEC-2000.
XX
PF 26-MAY-2000; 2000WO-US14842.
XX

PR 28-MAY-1999; 99US-0136503.
XX (SUGEN-) SUGEN INC.
XX
PI Plowman GD, Martinez R, Whyte D, Sudersanam S;
XX
DR WPI; 2001-032161/04.
DR P-PSDB; AAB65643.
XX
PT Nucleic acids encoding kinase polypeptides, useful for diagnosing and
PT treating immune-related diseases and disorders, cardiovascular disease,
PT neurodegenerative diseases and/or cancers -
XX
PS Disclosure; Fig 2; 310pp; English.
XX
CC The present sequence encodes a novel protein kinase. The nucleic acids
CC and the protein kinases they encode may be used in the treatment and
CC diagnosis of diseases associated with inappropriate kinase expression
CC such as immune-related diseases and disorders, cardiovascular disease,
CC neurodegenerative diseases and/or cancers. The nucleic acids and
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays. The kinase polypeptides may be used as antigens in the production
CC of antibodies of kinase expression and activity. Anti-kinase antibodies
CC and kinase antagonists may also be used to down regulate kinase
CC expression and activity. Diseases related to kinase expression and
CC activity include rheumatoid arthritis, atherosclerosis, autoimmune
CC disorders, complications of organ transplantation, myocardial infarction,
CC immune disorders, cardiomyopathies, strokes, renal failure,
CC oxidative-stress related disorders, chronic inflammatory bowel disease,
CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,
CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and
CC reproductive disorders.
XX
SQ Sequence 2615 BP; 742 A; 580 C; 585 G; 708 T; 0 other;

Query Match 59.2%; Score 559.2; DB 22; Length 2615;
Best Local Similarity 81.8%; Pred. No. 9.7e-141;
Matches 658; Conservative 0; Mismatches 143; Indels 3; Gaps 1;

QY 135 GAAAATAGCTAAGAGAAATACGATGTTGAAGCAATTAACCATCCAAATCTTGGA 194
DB 1 GAAAATAGCTCTCGGGAAATCCGATGCTGAAG---TTGAAACACCCAAACCTCGTGA 57

QY 195 CCTCATCGAGTGTTCAGGAGAAAAGGAAATGCAATTTAGTTTGAATACTGTGATCA 254
DB 58 CCTCATCGAGTGTTCAGAGAAAGAGAAAGATGCACTCTAGTTTGTGACTGTGATCA 117

QY 255 TACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTGTCTGATGGAGTCAAAAG 314
DB 118 CACACTGTTAAACGAGCTGGAGAGAAACCCAAACGGAGTTTCTGATGGAGTAAAG 177

QY 315 CGTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGTCATATACATAACTGTATTCAG 374
DB 178 TGTGCTATGSCAAACCTTCAAGCCCTTAACCTTCTGTCAAGCACAAATTGTATTCAG 237

QY 375 AGATATAAACCTGAAATATTCTAATAACTAAGCAAGGAATAATCAAGATTTGTGACTT 434
DB 238 GGATGTAAACCTGAAACACATCCTAATAACCAAGCAAGGATGATAAGATTTGTGACTT 297

QY 435 CGGGTTTGCAAAATTTCTGATTCAGGAGATGCCTACACCGATTATGTAGCTACGAGATG 494
DB 298 TGGATTTGCACGAATTTCTAATTCAGGAGACGCCTACACAGACTATGTTGCCACAGGTG 357

QY 495 GTACCGAGCTCCTGAACCTTCTTGTGGGAGATACCTCAGTATGGTTCTTTCAGTCGATATG 554
DB 358 GTACCGAGCCCGAACTTCTCGTGGGAGACACGAAGTACGGTTCCTCTGTAGACGTGTG 417

QY 555 GGCTATTGTTGTTTTCAGAGACTCCTGACAGGCCAGCCACTGTGGCTGGAAATC 614
DB 418 GGCCGTCCGCTGTGTTTTCAGAGACTCCTGACGGGTGACGCCACTCTGGCCGGGAAATC 477

QY 615 AGATGTGACCAACTTTTATCTGATAATCAGAACACTAGGAAATTAATCCCAAGACATCA 674
DB |||||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||

Dd 478 CGACGTGGACAGCTTTACCTGATCATCAGGACGTTGGGAAGCTGATTCGAAGACACCA 537
Qy 675 ATCAATCTTTAAAGTAACGGGTTTTTCCATGGCATCAGTATACCTGAGCCAGAAGACAT 734
Dd 538 GTCTATCTTTAGGAGTAACAGATTTTCCGGCGCATCAGCATACCTGACAGAGGACAT 597
Qy 735 GGAACCTCTTGAGGAAAGTTCTCAGATGTTTCATCCTGTGGCTCTGAACTTCATGAAGGG 794
Dd 598 GGAGACTCTTGGAAGAAAATTCTCAATGTTTCAGCCTGTGGCTTTAAGTTTCATGAAGGG 657
Qy 795 GTGTCTGAAGATGAATCCAGATGACAGATTAACTGTTCCTCCAACTCCCTGGAGAGCTCCTA 854
Dd 658 ATGCCCTGAAGATGAATCCTGATGAGAGGCTGACCTGTGCCCCAGCTGCTGGACAGTGCCTA 717
Qy 855 CTTTGATCTCTTTTCAAGAGGCGCAATTAAGAAAGAACGACGTAATGAGGAAGAAGACAG 914
Dd 718 CTTTGAGTCTTTTCAAGAGGATCAATGAATGAAGAAAGAGCCCGCAGTGAGGGGAGAGCCG 777
Qy 915 AAGACGCCAACAGGTACTTCCCGCT 938
Dd 778 AAGGCGCCAGCAGATCAACTGCT 801

RESULT 11
ABX34679
ID ABX34679 standard; cDNA; 1281 BP.

AC ABX34679;
XX
DT 13-FEB-2003 (first entry;
DE
XE Human mddt cDNA SEQ ID 240.
XX

KW MDDT; human; disease detection and treatment molecule polypeptide;
KW anti-inflammatory; immunosuppressive; osteopathic; cytostatic; anti-HIV;
KW haemostatic; nephrotropic; antianaemic; antipsoriatic; hepatotropic;
KW gene therapy; protein replacement therapy; cell proliferative disorder;
KW cancer; adenocarcinoma; leukaemia; lymphoma; melanoma; myeloma; sarcoma;
KW anaemia; Crohn's disease; acquired immunodeficiency syndrome; AIDS;
KW Goodpasture's syndrome; inflammation; osteoporosis; thrombocytopaenia;
KW psoriasis; hepatitis; gene; ss.

XX Homo sapiens.
OS
XX
PN WO200279449-A2.
XX

PD 10-OCT-2002.

XX 27-MAR-2002; 2002WO-US09944.

XX 28-MAR-2001; 2001US-279619P.

PR 29-MAR-2001; 2001US-280067P.

PR 29-MAR-2001; 2001US-280068P.

PR 16-MAY-2001; 2001US-291280P.

PR 17-MAY-2001; 2001US-291829P.

PR 17-MAY-2001; 2001US-291849P.

PR 19-JUN-2001; 2001US-299428P.

PR 20-JUN-2001; 2001US-299776P.

PR 20-JUN-2001; 2001US-300001P.

XX (INCY-) INCYTE GENOMICS INC.

XX Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;

PI Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshey SR;

PI Daugherty SC, Dam TC, Jiu TF, Nguyen DA, Kleefeld Y, Gerstin EH;

PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;

PI Flores V, Marwaha R, Lo A, Jan RY, Urashka ME;

PT (e.g. leukemia or sarcoma), anemia, Crohn's disease, AIDS, osteoporosis
PT or hepatitis
XX
PS Claim 1; SEQ ID NO 240; 339pp + Sequence Listing; English.
XX

CC This invention describes a novel disease detection and treatment molecule
CC polypeptide (MDDT) which has anti-inflammatory, immunosuppressive,
CC osteopathic, cytostatic, anti-HIV, haemostatic, nephrotropic,
CC antianaemic, antipsoriatic and hepatotropic activity. The polynucleotides
CC and the polypeptides of the invention can be used for gene therapy,
CC protein replacement therapy and are useful for treating a variety of
CC diseases or conditions. These polypeptides or polynucleotides are
CC particularly useful for diagnosing, treating or preventing cell
CC proliferative disorders (e.g. cancers including adenocarcinoma,
CC leukaemia, lymphoma, melanoma, myeloma or sarcoma), anaemia, Crohn's
CC disease, acquired immunodeficiency syndrome (AIDS), Goodpasture's
CC syndromes, inflammation, osteoporosis, thrombocytopaenia, psoriasis or
CC hepatitis. ABX34440-ABX34835 encode the MDDT polypeptides represented in
CC ABU11450-ABU11845, described in the disclosure of the invention.
CC NOTE: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 1281 BP; 428 A; 191 C; 317 G; 345 T; 0 other;

Query Match 57.5%; Score 543.2; DB 25; Length 1281;
Best Local Similarity 99.5%; Pred. No. 1.5e-136;
Matches 545; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATGGAAAAGTATGAAAATTTAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATCAAA 60
Dd 704 ATGGAAAAGTATGAAAATTTAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATCAAA 763

Qy 61 TGCAGAAACAAAACCTCTCGACAACTAGTAGCTGTTAAAAAATTTGTGGAATCTGAAGAT 120
Dd 764 TGCAGAAACAAAACCTCTCGACAACTAGTAGCTGTTAAAAAATTTGTGGAATCTGAAGAT 823

Qy 121 GATCCTGTTGTTAAGAAAATAGCACTAAGAGAAATACGTAATGTTGAAGCAATTAACAAT 180
Dd 824 GATCCTGTTGTTAAGAAAATAGCACTAAGAGAAATACGTAATGTTGAAGCAATTAACAAT 883

Qy 181 CCAATCTTGTGAACCTCATCGAGGTGTTTCCAGAGAAAAGAAAATGCAATTTAGTTT 240
Dd 884 CCAATCTTGTGAACCTCATCGAGGTGTTTCCAGAGAAAAGAAAATGCAATTTAGTTT 943

Qy 241 GAATACCTGTGATCATACACTTTTAAATGAGCTGGAAGAAAACCCCAATGGAGTTGCTGAT 300
Dd 944 GAATACCTGTGATCATACACTTTTAAATGGGCTGGAAGAAAACCCCAATGGAGTTGCTGAT 1003

Qy 301 GGAGTGATCAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGTCATATACAT 360
Dd 1004 GGAGTGATCAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGTCATATACAT 1063

Qy 361 AACTGTATTCACAGAGATATAAAACCTGAAAATATTCTAATAACTAAGCAAGGAATAATC 420
Dd 1064 AACTGTATTCACAGAGATATAAAACCTGAAAATATTCTAATAACTAAGCAAGGAATAATC 1123

Qy 421 AAGATTGTTGACTTCGGGTTTGCAAAAATTTCTGATTCAGGAGATGCCTACACCGATTAT 480
Dd 1124 AAGATTGTTGACTTCGGGTTTGCAAAAATTTCTGATTCAGGAGATGCCTACACCGATTAT 1183

Qy 481 GTAGCTACGAGATGGTACCGAGCTCTGAACTTCTTGTGGGAGATACCTCAGTATGTTCT 540
Dd 1184 GCAGCTACGAGATGGTACCGAGCTCTGAACTTCTTGTGGGAGATACCTCAGTATGTTCT 1243

Qy 541 TCAGTCGA 548
Dd 1244 TCAGTCAA 1251

RESULT 12
AAD03818
ID AAD03818 standard; cDNA; 1819 BP.

XX AAD03818;
XX 19-JUN-2001 (first entry)
XX Human kinase cDNA #7.
XX
XX Human; kinase; gene therapy; bioreactor; mental disorder;
XX biological disorder; polymorphism; ss.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX 5'UTR 1..137
XX /*tag= a
XX CDS 138..701
XX /*tag= b
XX /product= "Human kinase #1"
XX 3'UTR 702..1819
XX /*tag= c
XX misc_difference 1684
XX /*tag= d
XX /note= "polymorphism site; given as r in the sequence"
XX
XX WO200123579-A1.
XX
XX 05-APR-2001.
XX
XX 27-SEP-2000; 2000WO-US26621.
XX
XX 28-SEP-1999; 99US-015651i.
XX (LEXI-) LEXICON GENETICS INC.
XX
XX Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
XX WPI; 2001-266166/27.
XX P-PSDB; AAE00490.
XX
XX New isolated human kinase polynucleotide useful for generating
XX antibodies, as reagents in diagnostic assays and for screening for
XX compounds useful for treating mental, biological or medical diseases
XX
XX Disclosure; Page 34-35; 38pp; English.
XX
XX The present sequence encodes a novel human protein (NHP) which
XX is a human kinase. A polymorphism was identified in the 3' UTR
XX of the present sequence. The human kinases share structural
XX similarity with animal kinases, more particularly serine or
XX threonine protein kinases. Human kinase cDNA is useful for the
XX detection of mutant human kinase for the diagnosis of disease,
XX and also as a therapeutic. It is useful for screening drugs
XX effective in the treatment of symptomatic or phenotypic
XX manifestations perturbing the normal function of NHP in the
XX body. The NHP nucleotide sequences are useful for generation of
XX antibodies, as reagents in diagnostic assays, for the
XX identification of other cellular gene products related to human
XX kinases, and as reagents in assays for screening compounds that
XX are useful for treating mental, biological or medical disorders.
XX NHP oligonucleotides are used as probes. The labelled NHP probes
XX are useful for screening human genomic library for identifying
XX polymorphisms and as primers in amplification assays to detect
XX mutations within the exons, introns and splice sites that can
XX be used in diagnostics and pharmacogenomics. Nucleotide construct
XX encoding NHP products are used to genetically engineer cells
XX in vivo that functions as bioreactors in the body delivering a
XX continuous supply of NHP to the body. Nucleotide constructs
XX encoding functional NHPs are used in gene therapy for the
XX modulation of NHP expression.
XX
XX Sequence 1819 BP; 539 A; 368 C; 406 G; 505 T; 1 other;
SQ Query Match 53.7%; Score 507; DB 22; Length 1819;

Best Local Similarity 69.8%; Pred. No. 1.le-126;
Matches 929; Conservative 0; Mismatches 0; Indels 402; Gaps 2;
QY 1 ATGGAAGATATGAAAAATTAGCTAAGACTGGAGAAGGGTCTTATGGGTTGTATCAAA 60
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
138 ATGGAAGATATGAAAAATTAGCTAAGACTGGAGAAGGGTCTTATGGGTTGTATCAAA 197
QY 61 TGCAGAAACAAAACCTCTGGACAAGTAGTAGCTGTTAAAAAATTTGTGGAATCTGAAGAT 120
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
198 TGCAGAAACAAAACCTCTGGACAAGTAGTAGCTGTTAAAAAATTTGTGGAATCTGAAGAT 257
QY 121 GATCCTGTTGTTAAGAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACACAT 180
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
258 GATCCTGTTGTTAAGAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACACAT 317
QY 181 CCAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAGGAAAATGCATTTAGTTTTT 240
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
318 CCAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAGGAAAATGCATTTAGTTTTT 377
QY 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGGAGTTGCTGAT 300
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
378 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGGAGTTGCTGAT 437
QY 301 GGAGTGATCAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGTCATATACAT 360
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
438 GGAGTGATCAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGTCATATACAT 497
QY 361 AACTGTATTCACAGAGATATAAAACCTGAAAATATTCTAATAACAAAGGAATAATC 420
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
498 AACTGTATTCACAGAGATATAAAACCTGAAAATATTCTAATAACAAAGGAATAATC 557
QY 421 AAGATTTGTGACTTCGGGTTTGCAAAAATTCTG ----- 453
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
558 AAGATTTGTGACTTCGGGTTTGCAAAAATTCTGAGTTGGACTTCATCTTTCTCTGGTGCC 617
QY 454 ----- 453
DB TCCTTGATTGGCTTAAATAGTTGACCTTCTGAATTTCTTTCTGCCAATTCAGAGATTTT 677
QY 454 ----- 453
DB CTCCTGGCTTGGATCCATTGCTGACACAGTGTTTCACCAATGGGGCCAGGCTCATCTCGA 737
QY 454 ----- 453
DB AGCCACCGTGCCCGCCAGATTTTTTCAAAACAATACTACTGAGAGCTCAAGATTGTTT 857
QY 454 -----ATCCAGGAGATGCCTACACC 474
DB TTAGTGGGAACACAATTTCGAACAAATTTCTTGAGAACCGCATCCAGGAGATGCCTACACC 917
QY 475 GATTATGTAGCTACGAGATGGTACCGAGCTCCTGAACTTCTTGTGGGAGATACCTAGTAT 534
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
918 GATTATGTAGCTACGAGATGGTACCGAGCTCCTGAACTTCTTGTGGGAGATACCTAGTAT 977
QY 535 GGTTCTTCAGTCGATATATGGGCTATTGGTTGTGTTTTTGCAGAGCTCCTGACAGGCCAG 594
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
978 GGTTCTTCAGTCGATATATGGGCTATTGGTTGTGTTTTTGCAGAGCTCCTGACAGGCCAG 1037
QY 595 CCACTGTGGCTGAAAATCAGATGTGGACCAACTTTTATCTGATAATCAGAACACT--- 650
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1038 CCACTGTGGCTGAAAATCAGATGTGGACCAACTTTTATCTGATAATCAGAACACTAGTA 1097
QY 651 ----- 650
DB GAGACGGGGTTTCGCCATGTTTGACCAGGCTGGTCTCGAACTCTTGACGTCAAGTGATCCA 1157
QY 651 -----AGGAAATTAATCCCAAGACATCAATCA 678
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 1158 CCTGCCGTAGCCTCTCAAAAGTCTGGAAATTACAGGAAATTAATCCCAAGACATCAATCA 1217
Qy 679 ATCTTTAAAGTAACGGGTTTTTCCATGGCATCAGTATACCTGAGCCAGAGACATGGAA 738
Db 1218 ATCTTTAAAGTAACGGGTTTTTCCATGGCATCAGTATACCTGAGCCAGAGACATGGAA 1277
Qy 739 ACTCTTGAGGAAAGTTCTCAGATGTTTCATCCTCTGGCTCTGAACCTTCATGAAGGGGTGT 798
Db 1278 ACTCTTGAGGAAAGTTCTCAGATGTTTCATCCTCTGGCTCTGAACCTTCATGAAGGGGTGT 1337
Qy 799 CTGAAGATGAATCCAGATGACAGATTAACTGTTCCCAACTCCTGGAGAGCTCCTACTTT 858
Db 1338 CTGAAGATGAATCCAGATGACAGATTAACTGTTCCCAACTCCTGGAGAGCTCCTACTTT 1397
Qy 859 GATTCTTTCAAGAGGCCCAATTAAGAAAGACGTAATGAAGGAGAAACAGAGA 918
Db 1398 GATTCTTTCAAGAGGCCCAATTAAGAAAGACGTAATGAAGGAGAAACAGAGA 1457
Qy 919 CGCCAACAGGT 929
Db 1458 CGCCAACAGGT 1468

RESULT 13
ABZ77165
ID ABZ77165 standard; cDNA; 1086 BP.
XX
AC ABZ77165;
XX
DT 07-MAY-2003 (first entry)
XX
DE Human protein kinase encoding cDNA SEQ ID NO:79.
XX
KW Human; protein kinase; enzyme; antiasthmatic; antiinflammatory;
KW antidiabetic; antiparkinsonian; antimigraine; cardiant; cytostatic;
KW immunosuppressive; vulnery; gene therapy; COPD; asthma; migraine;
KW chronic obstructive pulmonary disease; non-insulin dependent diabetes;
KW Parkinson's disease; myocardial infarction; inflammatory bowel disease;
KW autoimmune disorder; allograft rejection; graft versus host disease;
KW cancer; leukaemia; wound granulation; gene; ss.
XX
OS Homo sapiens.

Key Location/Qualifiers
CDS 1..1086
/*tag= a
/partial
/product= "protein kinase"
/note= "no start or stop codons given"

WO2003000901-A2.

03-JAN-2003.

24-JUN-2002; 2002WO-IB02358.

26-JUN-2001; 2001US-301098P.

06-NOV-2001; 2001US-332870P.

(DECO-) DECODE GENETICS EHF.

Martinez RAM, Sigurdsson GT;

WPI; 2003-201429/19.

P-PSDB; ABP96087.

XX New protein kinase genes and polypeptides, useful for diagnosing
PT diseases associated with a protein kinase, or in gene therapy for
PT treating e.g. Parkinson's disease, migraine, myocardial infarction,
PT allograft rejection or cancers -

Claim 1; Page 89; 258pp; English.

CC ABZ77126 to ABZ77165 encode the human protein kinases given in ABP96048
CC to ABP96087. The protein kinases have antiasthmatic, antiinflammatory,
CC antidiabetic, antiparkinsonian, antimigraine, cardiant, cytostatic,
CC immunosuppressive and vulnery activities, and can be used in gene
CC therapy. A protein kinase therapeutic agent from the present invention,
CC particularly a protein kinase gene agonist or antagonist, can be used
CC for treating a disease or condition associated with a protein kinase in
CC an individual. These diseases include chronic obstructive pulmonary
CC diseases (COPD), asthma, non-insulin dependent diabetes, Parkinson's
CC disease, migraine, myocardial infarction, inflammatory bowel disease,
CC autoimmune disorders (e.g. allograft rejection or graft vs. host
CC disease), cancers (e.g. leukaemias) or wound granulation.

XX
SQ Sequence 1086 BP; 356 A; 227 C; 210 G; 293 T; 0 other;

Query Match 53.0%; Score 501; DB 25; Length 1086;
Best Local Similarity 100.0%; Pred. No. 3.6e-125;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 292 GTTGCTGATGGAGTGATCAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGT 351
Db 283 GTTGCTGATGGAGTGATCAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGT 342
Qy 352 CATATACATAACTGTATTCACAGAGATATAAAACCTGAAATATTTCTAATACTAAGCAA 411
Db 343 CATATACATAACTGTATTCACAGAGATATAAAACCTGAAATATTTCTAATACTAAGCAA 402
Qy 412 GGAATAATCAAGATTTGTGACTTCGGGTTTGCACAAATCTGATTCAGGAGATGCCTAC 471
Db 403 GGAATAATCAAGATTTGTGACTTCGGGTTTGCACAAATCTGATTCAGGAGATGCCTAC 462
Qy 472 ACCGATTATGTAGTACGAGATGGTACCGAGCTCCTGAACCTTCTTGGGAGATACCTAG 531
Db 463 ACCGATTATGTAGTACGAGATGGTACCGAGCTCCTGAACCTTCTTGGGAGATACCTAG 522
Qy 532 TATGGTTCTTCAGTCGATATATGGGCTATTGGTTGTTGTTTTCAGAGCTCCTGACAGGC 591
Db 523 TATGGTTCTTCAGTCGATATATGGGCTATTGGTTGTTGTTTTCAGAGCTCCTGACAGGC 582
Qy 592 CAGCCACTGTGGCTGGAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACTA 651
Db 583 CAGCCACTGTGGCTGGAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACTA 642
Qy 652 GGAATAATTAATCCCAAGACATCAATCAATCTTTAAAGTAACGGGTTTTTCCATGGCATC 711
Db 643 GGAATAATTAATCCCAAGACATCAATCAATCTTTAAAGTAACGGGTTTTTCCATGGCATC 702
Qy 712 AGTATACCTGAGCCAGAGACATGGAACACTCTTGAGGAAAGTTCTCAGATGTTTCATCCT 771
Db 703 AGTATACCTGAGCCAGAGACATGGAACACTCTTGAGGAAAGTTCTCAGATGTTTCATCCT 762
Qy 772 GTGGCTCTGAACTTCATGAAG 792
Db 763 GTGGCTCTGAACTTCATGAAG 783

RESULT 14
AAD03812
ID AAD03812 standard; cDNA; 561 BP.
XX
AC AAD03812;
XX 19-JUN-2001 (first entry)
XX Human kinase cDNA #1.

XX Human; kinase; gene therapy; bioreactor; mental disorder;
KW biological disorder; ss.
XX Homo sapiens.

XX Key Location/Qualifiers
FT CDS 1..561


```
FT      /*tag= a
FT      /product= "Human kinase #1"
FT      /notes= "The coding region does not include stop codon"
XX      /partial
PN      WC200123579-A1.
XX
XX      PD      05-APR-2001.
XX      PF      27-SEP-2000; 2000WO-US26621.
XX      PR      29-SEP-1999; 99US-0156511.
XX      PA      (LEXI-) LEXICON GENETICS INC.
XX
XX      P:      Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
XX      DR      WPI; 2001-266166/27.
XX      DR      P-PSDB; AAEC0490.
XX
XX      PT      New isolated human kinase polynucleotide useful for generating
XX      PT      antibodies, as reagents in diagnostic assays and for screening for
XX      PT      compounds useful for treating mental, biological or medical diseases
XX
XX      PS      Disclosure; Page 27; 38pp; English.
XX
XX      CC      The present sequence is a cDNA encoding novel human protein
XX      CC      (NHP) known as human kinase. The human kinases share structural
XX      CC      similarity with animal kinases, more particularly serine or
XX      CC      threonine protein kinases. Human kinase cDNA is useful for the
XX      CC      detection of mutant human kinase for the diagnosis of disease,
XX      CC      and also as a therapeutic. It is useful for screening drugs
XX      CC      effective in the treatment of symptomatic or phenotypic
XX      CC      manifestations perturbing the normal function of NHP in the
XX      CC      body. The NHP nucleotide sequences are useful for generation of
XX      CC      antibodies, as reagents in diagnostic assays, for the
XX      CC      identification of other cellular gene products related to human
XX      CC      kinases, and as reagents in assays for screening compounds that
XX      CC      are useful for treating mental, biological or medical disorders.
XX      CC      NHP oligonucleotides are used as probes. The labelled NHP probes
XX      CC      are useful for screening human genomic library for identifying
XX      CC      polymorphisms and as primers in amplification assays to detect
XX      CC      mutations within the exons, introns and splice sites that can
XX      CC      be used in diagnostics and pharmacogenomics. Nucleotide construct
XX      CC      encoding NHP products are used to genetically engineer cells
XX      CC      in vivo that functions as bioreactors in the body delivering a
XX      CC      continuous supply of NHP to the body. Nucleotide constructs
XX      CC      encoding functional NHPs are used in gene therapy for the
XX      CC      modulation of NHP expression.
XX
XX      SQ      Sequence 561 BP; 189 A; 87 C; 112 G; 173 T; 3 other;
XX
XX      Query Match      48.1%; Score 454.4; DB 22; Length 561;
XX      Best Local Similarity 99.8%; Pred. No. 1.1e-112;
XX      Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY      1      ATGGAAGAAGTATGAAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGGTGTATTCAA 60
Dh      1      ATGGAAGAAGTATGAAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGGTGTATTCAA 60
QY      61      TGCAGAAACAAACCTCTGGACAAGTAGTAGCTGTAAAGAAATTTTGGGAATCTGAAGAT 120
Dh      61      TGCAGAAACAAACCTCTGGACAAGTAGTAGCTGTAAAGAAATTTTGGGAATCTGAAGAT 120
QY      121      GATCCTGTGTGAAGAAATAGCACTAAGAGAAATACCTATGTTGAAGCAATTAACAT 180
Dh      121      GATCCTGTGTGAAGAAATAGCACTAAGAGAAATACCTATGTTGAAGCAATTAACAT 180
QY      181      CCAATCTGTGAACCTCATCGAGGTGTTTCAGGAGAAAGGAAATGCAATTTT 240
Dh      181      CCAATCTGTGAACCTCATCGAGGTGTTTCAGGAGAAAGGAAATGCAATTTT 240
QY      241      GAATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT 300
```

```
Db      241      GAATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT 300
QY      301      GGAGTGATCAAAAGCGTATTATGGCAACACACCTTCAAGCTCTTAATTTCTGTCATATACAT 360
Dh      301      GGAGTGATCAAAAGCGTATTATGGCAACACACCTTCAAGCTCTTAATTTCTGTCATATACAT 360
QY      361      AACTGTATTACAGAGATATAAAACCTGAAATATTTCTAATAACTAGCAAGGAATAATC 420
Dh      361      AACTGTATTACAGAGATATAAAACCTGAAATATTTCTAATAACTAGCAAGGAATAATC 420
QY      421      AAGATTGTGACTTCGGGTTTGCACAAATTTCTGATT 456
Dh      421      AAGATTGTGACTTCGGGTTTGCACAAATTTCTGATT 456

RESULT 15
AAD03815
ID      AAD03815 standard; cDNA; 594 BP.
XX
XX      AC      AAD03815;
XX      DT      19-JUN-2001 (first entry)
XX      DE      Human kinase cDNA #4.
XX      KW      Human; kinase; gene therapy; bioreactor; mental disorder;
XX      KW      biological disorder; ss.
XX      OS      Homo sapiens.
XX      FH      Key
XX      FT      CDS
XX      FT      Location/Qualifiers
XX      FT      1..594
XX      FT      /*tag= a
XX      FT      /product= "Human kinase #4"
XX      FT      /note= "The coding region does not include stop codon"
XX      FT      /partial
XX      PN      WC200123579-A1.
XX      PD      05-APR-2001.
XX      PF      27-SEP-2000; 2000WO-US26621.
XX      PR      29-SEP-1999; 99US-0156511.
XX      PA      (LEXI-) LEXICON GENETICS INC.
XX      P:      Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
XX      DR      WPI; 2001-266166/27.
XX      DR      P-PSDB; AAEC0493.
XX
XX      PT      New isolated human kinase polynucleotide useful for generating
XX      PT      antibodies, as reagents in diagnostic assays and for screening for
XX      PT      compounds useful for treating mental, biological or medical diseases
XX
XX      PS      Disclosure; Page 31; 38pp; English.
XX
XX      CC      The present sequence is a cDNA encoding novel human protein
XX      CC      (NHP) known as human kinase. The human kinases share structural
XX      CC      similarity with animal kinases, more particularly serine or
XX      CC      threonine protein kinases. Human kinase cDNA is useful for the
XX      CC      detection of mutant human kinase for the diagnosis of disease,
XX      CC      and also as a therapeutic. It is useful for screening drugs
XX      CC      effective in the treatment of symptomatic or phenotypic
XX      CC      manifestations perturbing the normal function of NHP in the
XX      CC      body. The NHP nucleotide sequences are useful for generation of
XX      CC      antibodies, as reagents in diagnostic assays, for the
XX      CC      identification of other cellular gene products related to human
XX      CC      kinases, and as reagents in assays for screening compounds that
XX      CC      are useful for treating mental, biological or medical disorders.
XX      CC      NHP oligonucleotides are used as probes. The labelled NHP probes
```

CC are useful for screening human genomic library for identifying
CC polymorphisms and as primers in amplification assays to detect
CC mutations within the exons, introns and splice sites that can
CC be used in diagnostics and pharmacogenomics. Nucleotide construct
CC encoding NHP products are used to genetically engineer cells
CC in vivo that functions as bioreactors in the body delivering a
CC continuous supply of NHP to the body. Nucleotide constructs
CC encoding functional NHPs are used in gene therapy for the
CC modulation of NHP expression.

XX

SQ Sequence 594 BP; 203 A; 94 C; 115 G; 182 T; 0 other;

Query Match 48.1%; Score 454.4; DB 22; Length 594;
Best Local Similarity 99.8%; Pred. No. 1.1e-112;
Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	ATGGAAAGTATGAAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGGTTGTATTCAAA	60
Db	1	ATGGAAAGTATGAAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGGTTGTATTCAAA	60
QY	61	TGCAGAAACAAACCTCTGGACAAGTAGTAGCTGTTAAAAAATTTGTGGAATCTGAAGAT	120
Db	61	TGCAGAAACAAACCTCTGGACAAGTAGTAGCTGTTAAAAAATTTGTGGAATCTGAAGAT	120
QY	121	GATCCTGTGTTAAGAAAAATAGCACTAAGAGAAATACCTATGTTGAAGCAATTAACAAT	180
Db	121	GATCCTGTGTTAAGAAAAATAGCACTAAGAGAAATACCTATGTTGAAGCAATTAACAAT	180
QY	181	CCAAATCTGTGAACCTCATCGAGGTGTTTAAATGAGCTGGAAGAAACCCAAATGCAATTTAGTTT	240
Db	181	CCAAATCTGTGAACCTCATCGAGGTGTTTAAATGAGCTGGAAGAAACCCAAATGCAATTTAGTTT	240
QY	241	GAATACCTGTGATCATACACCTTTAAATGAGCTGGAAGAAACCCAAATGCAATTTAGTTT	300
Db	241	GAATACCTGTGATCATACACCTTTAAATGAGCTGGAAGAAACCCAAATGCAATTTAGTTT	300
QY	301	GGAGTGATCAAAAGCGTATTATGCAAAACACTTCAAGCTCTTAATTTCTGCATATACAT	360
Db	301	GGAGTGATCAAAAGCGTATTATGCAAAACACTTCAAGCTCTTAATTTCTGCATATACAT	360
QY	361	AACTGTATTACAGAGATATAAAACCTGAAATATTCTAATAACTAAGCAAGGAATAATC	420
Db	361	AACTGTATTACAGAGATATAAAACCTGAAATATTCTAATAACTAAGCAAGGAATAATC	420
QY	421	AAGATTTGTGACTTCGGGTTTGCACAAATTCGATT	456
Db	421	AAGATTTGTGACTTCGGGTTTGCACAAATTCGATT	456

Search completed: November 15, 2003, 16:54:33
Job time : 301.918 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2003, 16:54:37 ; Search time 67.0921 Seconds
(without alignments)
6216.928 Million cell updates/sec

Title: US-09-671-050-11
Perfect score: 945
Sequence: 1 atggaaaagtatgaaaaatt.....aggtacttcggtcaaaagt 945

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1:39956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : issued_patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTJS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfilesi.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	287	30.4	1701	4 US-09-411-628-5	Sequence 5, Appli
2	287	30.4	3080	4 US-09-411-628-3	Sequence 3, Appli
3	171.8	18.2	903	2 US-08-874-347-9	Sequence 9, Appli
4	171.8	18.2	903	3 US-09-093-522-9	Sequence 9, Appli
5	163.4	17.3	1476	2 US-08-969-106-1	Sequence 1, Appli
6	163.4	17.3	1476	4 US-09-338-125-1	Sequence 1, Appli
7	163.4	17.3	1476	4 US-09-266-225D-13	Sequence 13, Appli
8	163.4	17.3	1635	4 US-09-417-197-112	Sequence 112, App
9	163.4	17.3	1635	4 US-09-417-197-114	Sequence 114, App
10	155.6	16.5	1050	4 US-09-220-132-3	Sequence 3, Appli
11	136.6	14.5	1825	4 US-09-620-312D-313	Sequence 313, App
12	134.6	14.2	1089	1 US-08-154-915-1	Sequence 1, Appli
13	134.6	14.2	1089	2 US-08-464-517-37	Sequence 37, Appli
14	134.6	14.2	1089	2 US-08-246-361A-37	Sequence 37, Appli
15	134.6	14.2	1089	3 US-08-463-772-37	Sequence 37, Appli
16	134.6	14.2	1089	5 PCT-US93-09945-1	Sequence 1, Appli
17	131	13.9	1002	1 US-08-463-090B-3	Sequence 3, Appli
18	117.2	12.4	1070	1 US-08-463-090B-5	Sequence 5, Appli
19	112	11.9	1308	4 US-09-801-861-4	Sequence 4, Appli
20	112	11.9	2203	4 US-09-801-861-1	Sequence 1, Appli
21	103	10.9	2747	2 US-08-874-347-1	Sequence 1, Appli
22	103	10.9	2747	3 US-09-093-522-1	Sequence 1, Appli
23	102	10.8	2467	4 US-09-206-344A-3	Sequence 3, Appli
24	98	10.4	1673	3 US-09-347-801-15	Sequence 15, Appli
25	97.8	10.3	1100	2 US-08-950-449A-18	Sequence 18, Appli
26	97.8	10.3	3774	2 US-08-950-449A-11	Sequence 11, Appli
27	97.8	10.3	3813	1 US-08-469-421-11	Sequence 11, Appli

28	97.8	10.3	3813	1 US-08-250-975-11	Sequence 11, Appli
29	97.8	10.3	3813	2 US-08-605-002A-11	Sequence 11, Appli
30	97.8	10.3	3813	4 US-08-943-353-11	Sequence 11, Appli
31	97.8	10.3	3813	5 PCT-US94-10529-11	Sequence 11, Appli
32	94.8	10.0	1272	4 US-09-206-344A-1	Sequence 1, Appli
33	94.8	10.0	1549	4 US-09-206-344A-13	Sequence 13, Appli
34	88.2	9.3	1381	2 US-08-950-449A-13	Sequence 13, Appli
35	88.2	9.3	1423	1 US-08-469-421-13	Sequence 13, Appli
36	88.2	9.3	1423	1 US-08-250-975-13	Sequence 13, Appli
37	88.2	9.3	1423	2 US-08-605-002A-13	Sequence 13, Appli
38	88.2	9.3	1423	4 US-08-943-353-13	Sequence 13, Appli
39	88.2	9.3	1423	5 PCT-US94-10529-13	Sequence 13, Appli
40	88.2	9.3	1539	3 US-09-286-904-1	Sequence 1, Appli
41	88.2	9.3	1539	4 US-09-640-101-1	Sequence 1, Appli
42	88.2	9.3	1821	4 US-09-417-197-64	Sequence 64, Appli
43	88.2	9.3	1824	4 US-09-417-197-46	Sequence 46, Appli
44	86.6	9.2	2624	1 US-08-032-382B-1	Sequence 1, Appli
45	85.4	9.0	1200	1 US-08-011-398B-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-411-628-5
; Sequence 5, Application US/09411628
; Patent No. 6428994
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; TITLE OF INVENTION: CDNA, GENOMIC, AND PREDICTED PROTEIN
; TITLE OF INVENTION: SEQUENCES OF LEARNING-INDUCED KINASES
; FILE REFERENCE: 13761-707
; CURRENT APPLICATION NUMBER: US/09/411,628
; CURRENT FILING DATE: 1999-10-01
; EARLIER APPLICATION NUMBER: US 60/102,906
; EARLIER FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1701
; TYPE: DNA
; ORGANISM: Orcytolagus cuniculus
US-09-411-628-5

Query Match	30.4%	Score	287;	DB	4;	Length	1701;
Best Local Similarity	59.1%	Pred. No.	1.9e-71;				
Matches	510;	Conservative	0;	Mismatches	350;	Indels	3;
Gaps	1;						
Qy	1	ATGGAAGATGATAAATAGCTAAGACTGGAGAGGGTCTTATGGGGTTGATTCAAA	60				
Db	1	ATGGAAGATGATAAATAGCTAAGACTGGAGAGGGTCTTATGGGGTTGATTCAAA	60				
Qy	61	TGCAGAAACAAACCTCTGGACAGTAGTAGCTGTTAAATAATTTGTGGAATCTGAAGAT	120				
Db	61	TGTAGGAATAAAGATAGTGGAGAAATTTGGCCATCAAGAAAGTTCTTAGAAAGTATGAT	120				
Qy	121	GATCCTGTTTAAAGAAATAGCACTAAGAGAAATACGTTATGTTGAAGCAATTAAACAT	180				
Db	121	GACAAATGGTTAAATAATTTGCTATCGGAGAAATCAAGTTACTAAAGCAACTGAGGCAT	180				
Qy	181	CCAAATCTTGAAACCTCATCGAGGTGTTTAAATAAGGAAATGCAATTAGTTTTT	240				
Db	181	GAAATTTGGTGAATCTGTTGGAGGTGTTGTAATAAATAAATAAATGATGTTACCTAGTCTTT	240				
Qy	241	GAATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT	300				
Db	241	GAATTTGTTGACCAACAGATTTCTTGATGACTTGGAACTCTTCCAAATGGACTAGATGAC	300				
Qy	301	GGAGTGATCAAAAGCGTATTATGCGAAACACTTCAAGCTCTTAATTTCTGTATATACAT	360				
Db	301	CAAGTAGTTCAAAAGTATTGTTTTCAGATTATTAAATGGAATGGATTTTGTTCACACTCAC	360				
Qy	361	AACTGTATTCACAGATATAAAACCTGAAATAATTTCTAATAACTAAGCAAGGAATAATC	420				

Db 361 AATATCATACATAGAGATATAAAGCCAGAGAAATATTTGGTCTCCAGTCTGGCGTTGC 420
Qy 421 AAGATTGTGACTCGGTTTGCACAAATCTG- - -ATCCAGGAGATGCCTACACCAT 477
Db 421 AAGATTGTGATTGGATTGACGGACACTGGCAGCTCCCGAGAGSTTTACACTGAT 480
Qy 478 TATGTAGCTACGAGATGGTACCGAGCTCCTGAACTTCTTGTTGGAGATACCTCAGTAJGT 537
Db 481 TATGTGGCAACTCGATGGTACAGAGCTCCAGAACTACTGGTTGGTGTGATGTCAGATGGC 540
Qy 538 TCTTCAGTCGATATATGGGCTATTGGTTGGTTTTCAGAGCTCCTGACAGGCCAGCCA 597
Db 541 AAGCTGTGGATGTGTGGCCATTGGTTGTCTGTAACCTGAAATGCTCAGTATGGGGAACCC 600
Qy 598 CTGTGGCCTGGAAATCAGATGTGGACCACTTTTATCTGATAATCAGAACACTAGGAAA 657
Db 601 CTGTTTCTCGAGACTCTGATATTGATCAGCTTTATCTTATTATGAGGTGTTTAGGTAAT 660
Qy 658 TTAATCCCAAGACATCAATCAATCTTTAAAGTAAAGGGTTTTCATGGCATCAGTATA 717
Db 661 CTAATCCCAAGACACAGGAGCTTTTATATAAAATCCTGTGTTTCTGGAGTAAGGTG 720
Qy 718 CCTGAGCCAGAGACATGGAACTCTTGAGGAAAGTTCTCAGATGTTTCATCCTGTGGCT 777
Db 721 CCTGAAATCAAGGAATCAGAACCTCTTGAAGACGCTATCCCAAGCTCTCAGAAAGTTG 780
Qy 778 CTGAACCTTCATGAAGGGTGTCTGAAGATGAATCCAGATGACAGATTAACCTGTCCCAA 837
Db 781 ATAGATTAGCAAGAAATGCTTACATGTTGACCCAGACAAAGGCCCTTCTGTGCTGAG 840
Qy 838 CTCCTGGAGAGCTCCTACTTTGA 860
Db 841 CTCCTACACCATGATTCTTTCA 863

RESULT 2

US-09-411-628-3
; Sequence 3, Application US/09411628
; Patent No. 6428994
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; TITLE OF INVENTION: CDNA, GENOMIC, AND PREDICTED PROTEIN
; TITLE OF INVENTION: SEQUENCES OF LEARNING-INDUCED KINASES
; FILE REFERENCE: 13761-707
; CURRENT APPLICATION NUMBER: US/09/411,628
; CURRENT FILING DATE: 1999-10-01
; EARLIER APPLICATION NUMBER: US 60/102,906
; EARLIER FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3080
; TYPE: DNA
; ORGANISM: Orcytolagus cuniculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (249)...(1949)
US-09-411-628-3

Query Match 30.4%; Score 287; DB 4; Length 3080;
Best Local Similarity 59.1%; Pred. No. 2.5e-71;
Matches 510; Conservative 0; Mismatches 350; Indels 3; Gaps 1;
Qy 1 ATGGAAGAAGTATGAAAAATTAGCTAAGACTGGAGAAGGGTCTTATGGGTTGTATTCAAA 60
Db 249 ATGGAAGAATATGAGAACCTTGGATTGGTTGGAGAAGGGAGTTATGGAATGGTATGAAG 308
Qy 61 TGCAGAACAAACCTCTGGACAAAGTAGTAGCTGTAAAAAATTTGTGGAATCTGAAGAT 120
Db 309 TGTAGGAATAAAGATAGTGGGAAGAATTGTGGCCATCAAGAAAGTTCTCAGAAAGTATGAT 368
Qy 121 GATCCTGTTGTTAAGAAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAACAT 180

Db 369 GACAAAAATGGTTAAAAAAATTGCTATGCGAGAAATCAAGTTACTAAAGCAACTGAGGCAT 428
Qy 181 CCAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAAGAAAAATGCATTTAGTTTTT 240
Db 429 GAAAAATTTGGTGAATCTGTTGGAGGTGTGTAAAAAAGAAAAAAGATGGTACCTAGTCTTT 488
Qy 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT 300
Db 489 GAATTTGTTGACCAACACGATTCTTGTGATGACTTGGAACTCTTCCAAATGGACTAGATGAC 548
Qy 301 GGAGTGATCAAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGTCAATATACAT 360
Db 549 CAAGTAGTTCAAAAGTATTGTTTTCAGATTATTAAATGGAATGGATTTTGTCCACAGTAC 608
Qy 361 AACTGTATTCACAGAGATATAAAACCTGAAAAATATTCTTAATAACAAAGCAAGGAATATC 420
Db 609 AATATCATACATAGAGATATAAAGCCAGAGAAATATATGTTCTCCAGTCTGGCGTTGTC 668
Qy 421 AAGATTGTGACTTCGGGTTTTCGACAAATTCG- - -ATTCCAGGAGATGCCTACACCGAT 477
Db 669 AAGTTATGCTGATTTTGGATTTCACGGACACTGGCAGCTCCCGAGAGAGTTTACACTGAT 728
Qy 478 TATGTAGCTACGAGATGGTACCGAGCTCCTGAACTTCTTGTGGGAGATACCTCAGTATGGT 537
Db 729 TATGTGGCAACTCGATGGTACAGAGCTCCAGAACTACTGTTGTTGATGTCAAGTATGGC 788
Qy 538 TCTTCAGTCGATATATGGGCTATTGGTTGTTGTTTTTTCAGAGACTCCTGACAGGCCAGCCA 597
Db 789 AAAGCTGTGGATGTGTGGGCCATTGGTTGTTGTTAACTGAAATGCTCATGAGGTGTTAGGTAAT 908
Qy 598 CTGTGGCCTGAAAAATCAGATGTGGACCAACTTTTATCTGATAATCAGAAACACTAGGAAA 657
Db 849 CTGTTTCTCGAGACTCTGATATTGATCAGCTTTTATATAAAATCCTGTGTTGCTGGAGTAAGGTTG 908
Qy 658 TTAATCCCAAGACATCAATCAATCTTTAAAAAGTAAAGGGTTTTCATGGCATCAGTATA 717
Db 909 CTAATTCAGACACACAGGAGCTTTTATATAAAATCCTGTGTTGCTGGAGTAAGGTTG 968
Qy 718 CCTGAGCCAGAGACATGGAACCTCTTGAGGAAAAAGTTCTCAGATGTTTCATCCTGTGGCT 777
Db 969 CCGAAATCAAGGAATCAGAACCTCTTGAAGAGCGCTATCCCAAGCTCTCAGAAAGTTG 1028
Qy 778 CTGAACCTTCATGAAGGGGTGCTGAAGATGAATCCAGATGACAGATTAACCTGTCCCAA 837
Db 1029 ATAGATTAGCAAGAAATGCTTACATGTTGACCCAGACAAAGGCCCTTCTGTGCTGAG 1088
Qy 838 CTCCTGGAGAGCTCCTACTTTGA 860
Db 1089 CTCCTACACCATGATTCTTTCA 1111

RESULT 3

US-08-874-347-9
; Sequence 9, Application US/08874347
; Patent No. 5863741
; GENERAL INFORMATION:
; APPLICANT: Limper, Andrew H.
; APPLICANT: Leof, Edward B.
; APPLICANT: Thomas, Charles F.
; APPLICANT: Gustafson, Michael P.
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
; TITLE OF INVENTION: CARINII
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C., P.A.
; STREET: 60 South Sixth Street, Suite 3300
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/874,347
FILING DATE: 13-JUN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ellinger, Mark S.
REGISTRATION NUMBER: 34,812
REFERENCE/DOCKET NUMBER: 07039/055001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-335-5070
TELEFAX: 612-288-9696
TELEX:

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 903 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: Coding Sequence

LOCATION: 1..900

OTHER INFORMATION:

US-08-974-347-9

Query Match 18.2%; Score 171.8; DB 2; Length 903;

Best Local Similarity 52.3%; Pred. No. 5e-39;

Matches 458; Conservative 0; Mismatches 402; Indels 15; Gaps 3;

QY 1 ATGGAAGAATGATGAAAAATTAGCTAAGACTGGAGAGGCTCTTATGGGTTGTATTCAA 60
DB 1 ATGGAGCAATATCAGAGGTAGAGAAAGATTGGAGAAAGCACTTATGGAGTTGTTATA 60
QY 61 TGCAGAAACAAAACCTCTGGACAAGTAGTAGCTGTTAAAAAATTTGTGGAATCTGAAG 120
DB 61 GCAAAGGATCTTGAAAGTGGTACAAATTGTAGCTCTTAAGAAAATCCGATTAGAAAG 120
QY 121 GATCCTGTTTAAAGAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAAACAT 180
DB 121 GATGAGGAGTTCCTAGTACAGCAATTCGTGAGATATCACTTTTGAAGAGATGCACA 180
QY 181 CCAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAGGAAAATGCATTAGTTT 240
DB 181 GATAATGTTGTAAGACTTTTGAATATAATTCATCAAGACTCACGTTTATATCTGTTT 240
QY 241 GAATCTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATG-----GAG 294
DB 241 GAATTTCTGTATCTTGATTTAAAAAAGTATATGAATAGTATTCCAAAGGACATGCTT 300
QY 295 GCTGATGGAGTGATCAAAAGCGTATTATGGCAACACCTTCAAGCTCTTAATTCTGTCA 354
DB 301 GGTGCAGAAATGATTAAAAAGTTTATGTCACAACTTGTATCAGGTGTTAAATATTGTC 360
QY 355 ATACATAACTGTATTTCACAGAGATATAAAACCTGAAATATTTCTAATAACTAAGCAAG 414
DB 361 TCTCATCGTATTCTTCATCGTGACTTGAAACCCACAAAATCTCTTATAGATCGAGAAG 420
QY 415 ATAATCAAGATTTGTGACTTCGGGTTTGCA---CAAATCTGATTCAGGAGATGCCTAC 471
DB 421 AATCTTAAATAGCAGATTTTGGGCTTGCAAGGGCGTTTGGTGTTCATTCGGTGGTTAT 480
QY 472 ACCGATTATGTAGCTACGAGATGTPACGAGCTCCTGAACTTCTTGTGGAGATACTCAG 531
DB 481 ACTCATGAAGTTGTTACACTTGGTATCGTGTCTCCAGAAAGTTCTTTTAGTGTGCGCAA 540
QY 532 TATGGTTCTTCAGTCGATATATGGGCTATTGGTTGTGTTTTCAGAGCTCCTGACAGGC 591
DB 541 TATGCAACAGCGCTTGATATATGGAGCAATTGGATGTTATTTTGCAGAAATGGGTACAAA 600
QY 592 CAGCCACTGTGGCCTGGAAAAATCAGATGTGGACCAACTTATCTGATATCAGAACAC 651

DB 601 AAGCCATTATTTCCAGGTGATTCTGAAATTGATGAAATATTTAGAATATTTAGAATATTA 660
QY 652 GGAATAATTAATCCCAAGACATCAATCAATCTTTAAAAAGTAACGGGTTTTTCCATGGCATC 711
DB 661 GGGACTCCAGATGAAAAATCTTGGCCTGCTGATTACATCTTATCCGATTTTAAAGGCACT 720
QY 712 AGTATACCTGAGCCAGAGACATGGAACACTCTTGAGGAAAAGTTCTCAGATGTTTCATCCT 771
DB 721 TTTCCAAAAATGTCACCA-----AAAAATCTTGAGAAATTAATTACAGAACTTGATAGT 774
QY 772 GTGGCTCTGAACCTTCATGAAGGGGTGCTCTGAAGATGAATCCAGATGACAGATTAACCTGT 831
DB 775 GATGGAATAGATTATTACAGAAATGCTCTAGATATTATCCTGCTGAAACGTATTAGCGCT 834
QY 832 TCCCAACTCCTGGAGAGCTCCTACTTTGATTCTTT 866
DB 835 AAAAAAGCTCTCGATCATCCTTATTTTGATGATT 869

RESULT 4

US-09-093-522-9

: Sequence 9, Application US/09093522

: Patent No. 6015700

: GENERAL INFORMATION:

: APPLICANT: Limper, Andrew H.

: APPLICANT: Leof, Edward B.

: APPLICANT: Thomas, Charles F.

: APPLICANT: Gustafson, Michael P.

: TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCOYSTIS

: TITLE OF INVENTION: CARINII

: NUMBER OF SEQUENCES: 26

: CORRESPONDENCE ADDRESS:

: ADDRESSEE: Fish & Richardson P.C., P.A.

: STREET: 60 South Sixth Street, Suite 3300

: CITY: Minneapolis

: STATE: MN

: COUNTRY: USA

: ZIP: 55402

: COMPUTER READABLE FORM:

: MEDIUM TYPE: Diskette

: COMPUTER: IBM Compatible

: OPERATING SYSTEM: DOS

: SOFTWARE: FastSeq for Windows Version 2.0

: CURRENT APPLICATION DATA:

: APPLICATION NUMBER: US/09/093,522

: FILING DATE: 08-JUN-1998

: CLASSIFICATION:

: PRIOR APPLICATION DATA:

: APPLICATION NUMBER: 08/874,347

: FILING DATE: 13-JUN-1997

: ATTORNEY/AGENT INFORMATION:

: NAME: Ellinger, Mark S.

: REGISTRATION NUMBER: 34,812

: REFERENCE/DOCKET NUMBER: 07039/055002

: TELECOMMUNICATION INFORMATION:

: TELEPHONE: 612-335-5070

: TELEFAX: 612-288-9696

: TELEX:

: INFORMATION FOR SEQ ID NO: 9:

: SEQUENCE CHARACTERISTICS:

: LENGTH: 903 base pairs

: TYPE: nucleic acid

: STRANDEDNESS: single

: TOPOLOGY: linear

: MOLECULE TYPE: cDNA

: NAME/KEY: Coding Sequence

: LOCATION: 1..900

: OTHER INFORMATION:

: US-09-093-522-9

Query Match 18.2%; Score 171.8; DB 3; Length 903;

```
Best Local Similarity 52.3%; Pred. No. 5e-39;
Matches 458; Conservative 0; Mismatches 402; Indels 15; Gaps 3;

Qy 1 ATGGAAGAAGTATGAAATTTAGCTAAGACTGGAGAAGGGCTTTATGGGTTGTATTCAAA 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 ATGGAGCAATATCAGAGGTTAGAGAAGATTGGAGAAGGAACCTTATGGAGTTGTTTATAAA 60

Qy 61 TGCAGAAACAAACCTCTGACAAAGTAGTAGCTTTAAAAAATTTGTGGAATCTGAAGAT 120
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 GCAAAGGATCTTGAAGTGGTACAATTTGAGCTTTAAGAAAATCCGATTAGAAGCAGAA 120

Qy 121 GATCCTGTTGTTAAGAAAATAGCACTAAGAGAATAACGTATGTTGAAGCAATTAACAACAT 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 GATGAGGGAGTTCCTAGTACAGCAATTCGTGAGATATCACTTTTGAAGAGATGCACAA 180

Qy 181 CCATATCTTGTGAACCTCATCGAGGTGTTTCAGAGAAAAGGAAAATGCAATTTAGTTTTT 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 GATAATGTTGTAAGACTTTTGAATATAATTCATCAAGAGTCACGTTTATACCTTGTGTTT 240

Qy 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAAAGAAACCCAAATG-----GAGTT 294
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 GAATTTCTTGATCTTGATTTTAAAAAGTATATGAATAGTATTTCCAAAGGACATGATGCTT 300

Qy 295 GCTGATGGAGTGATCAAAAGCGTATTATGGCAACACACTTCAAGCTCTTAATTTCTGTGAT 354
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 GGTGCAGAAATGATTAAAAAGTTTATGTCACAACTTGTATCAGGTGTTAAATATTGTCA 360

Qy 355 ATACATAACTGTATTCACAGAGATATAAAACCTGAAAATATTCTTAATAACTAAGCAAGGA 414
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 TCTCATCGTATTCCTCATCGTGACTTGAAACCCACAAAATCTTCTTATAGATCGAGAAGGA 420

Qy 415 ATAATCAAGATTTGTGACTTCGGGTTTGCA---CAAATTTCTGATTCAGGAGATGCCTAC 471
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 421 AATCTTAAATTAGCAGATTTTGGGCTTGCAAGGGCGTTTGGTGTTCCTATTCGTTGTTAT 480

Qy 472 ACCGATTATGTAGCTACGAGATGGTACCGAGCTCCTGAACTTCTTGTGGGAGATACCTAG 531
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 481 ACTCATGAAGTTGTTACACTTTGGTATCGTGTCCAGAAAGTTCTTTTAGTGGTGCACAA 540

Qy 532 TATGGTTCTTCAGTCGATATATGGGCTATTGGTTGTGTTTTTGCAGAGCTCCTGACAGGC 591
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 541 TATGCAACAGCGCTTGATATATGGAGCATTTGGATGTAATTTTGCAGAAATGGCTACAAA 600

Qy 592 CAGCCACTGTGGCCTGGAAAATCAGATGTGGACCAACTTTATCTGATAAATCAGAACACTA 651
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 601 AAGCCATTATTTCCAGGTGATTTCTGAAATTSATGAAATATTTAGAAATTTTAGAATATTA 660

Qy 652 GGAAAATTAATCCCAAGACATCAATCAATCTTTAAAAAGTAACGGGTTTTCCATGGCATC 711
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 661 GGGACTCCAGATGAAAATTTCTTGGCCTGGTATTACATCTTATCCGGATTTTAAAGGCAACT 720

Qy 712 AGTATACCTGAGCCAGACAGACATGGAACCTCTTGAGGAAAAGTTCTCAGATGTTTCATCCT 771
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 721 TTCCAAAATGGTCACCA-----AAAAATCTTGAGAAATTAATTACAGAACTTGAATAGT 774

Qy 772 GTGGCTCTGAACCTCATGAAGGGGTGCTGTAAGATGAATCCAGATGACAGATTAACCTGT 831
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 775 GATGGAATAGATTTTATTACAGAAATGCTCTTAGATATTATCCTGCTGACGTAATTAGCGCT 834

Qy 832 TCCCAACTCCTGGAGAGTCCCTACTTTTGAATCTTT 866
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 835 AAAAAGCTCTCGATCATCCTTATTTTGTATGATTT 869
```

RESULT 5
US-08-969-106-1
; Sequence 1, Application US/08969106
; Patent No. 5986055
; GENERAL INFORMATION:
; APPLICANT: Yang, M.
; APPLICANT: Nandabalan, K.
; APPLICANT: Schulz, V.
; TITLE OF INVENTION: CDK2 INTERACTIONS

```
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/969,106
; FILING DATE: 13-NOV-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7934-057
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1476 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 1...894
; OTHER INFORMATION:
; US-08-969-106-1
```

```
Query Match 17.3%; Score 163.4; DB 2; Length 1476;
Best Local Similarity 55.1%; Pred. No. 1.5e-36;
Matches 364; Conservative 0; Mismatches 291; Indels 6; Gaps 2;

Qy 1 ATGGAAGAAGTATGAAATTTAGCTAAGACTGGAGAAGGGTCTTATGGGTTGTATTCAAA 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 ATGGAGAACTTCCAAAAGGTGGAAAAGATCGGAGAGGGACGTCACGGAGTTGTGTACAAA 60

Qy 61 TGCAGAAACAAACCTCTGGAACAAGTAGACTAAGAGAAATACGTATGTTGAAGCAATTAACAACAT 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 GCCAGAAACAAGTTGACGGGAGAGGTGGTGGCGCTTAAGAAAATCCGCTGGACACTGAG 120

Qy 121 GATCCTGTTGTTAAGAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAACAT 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 ACTGAGGGTGTGCCAGTACTGCCATCCGAGAGATCTCTCTGCTTAAGGAGCTTAACCAT 180

Qy 181 CCATATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAGGAAAATGCAATTTAGTTTTT 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 CCTAATATTGCAAGCTGCTGGATGTCATTCAACAGAAAATAAACTCTACCTGGTTTTT 240

Qy 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGA---AAGAAAACCCAAATGGAGTTGCT 297
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 GAATTTCTGCACCAAGATCTCAAGAAATTCATGGATGCCTCTGCTCTCACTGGCATTCCT 300

Qy 298 GATGGAGTGAACAAAGCGTATTATGGCAACACACTTCAAGCTCTTAATTTCTGTATATA 357
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 CTTCCTCCTCATCAAGAGCTATCTGTTCCAGCTGCTCCAGGGCTAGCTTTCTGCCATTCT 360

Qy 358 CATAACTGTATTCACAGAGATATAAAACCTGAAAATATTCTAATAACTAAGCAAGGAATA 417
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 CATCGGGTCTCCACCGAGACCTTAACCTCAGAAATCTGCTTATTAACACAGAGGGGGGCC 420

Qy 418 ATCAAGATTTGTGACTTCGGGTTTGCACAAATTTCTGA---TTCAGGAGATGCCTACACC 474
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 421 ATCAAGCTACGAGACTTTGGACTAGCCAGAGCTTTTGGAGTCCCTGTTCTGCTACTTACACC 480
```

QY 475 GATTATGTAGCTACGAGATGGTACCGAGCTCCTGAACTTCTTGGGAGATACCTCAGTAT 534
| | | | |
Db 481 CATGAGGTGGTGACCTGTGGTACCGAGCTCCTGAAATCCTCTGGGCTCGAAATATTAT 540
| | | | |
QY 535 GGTCTTCAGTCGATATATGGCTATTGGTTGTGTTTTCAGAGCTCCTGACAGGCCAG 594
| | | | |
Db 541 TCCACAGCTGTGGACATCTGGAGCCTGGGCTGCATCTTCTGAGATGGTGAATCTGCGCG 600
| | | | |
QY 595 CCACTGTGGCCTGGAAATCAGATGTGGACCAACTTATCTGATAATCAGAACACTAGSA 654
| | | | |
Db 601 GCCCTGTCCCTGGAGATTCTGAGATTGACCAGCTCTTCCGGATCTTTCGGACTCTGGGG 660
| | | | |
QY 655 A 655
|
Db 661 A 661

RESULT 6
US-09-338-125-1
; Sequence 1, Application US/09338125
; Patent No. 6521412
; GENERAL INFORMATION:
; APPLICANT: Yang, M.
; APPLICANT: Nandabalan, K.
; APPLICANT: Schulz, V.
; TITLE OF INVENTION: CDK2 INTERACTIONS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/338,125
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/969,106
; FILING DATE: 13-NOV-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7934-057
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1476 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 1...894
; OTHER INFORMATION:
US-09-338-125-1

Query Match 17.3%; Score 163.4; DB 4; Length 1476;
Best Local Similarity 55.1%; Pred. No. 1.5e-36;
Matches 364; Conservative 0; Mismatches 291; Indels 6; Gaps 2;
QY 1 ATGGAAAAGTATGAAAAAATTAGCTAAGACTGGAGAGGGCTCTTATCGGGTGTATTCAAA 60

Db 1 ATGGAGAACTTCCAAAAGGTGAAAAGATCGGAGAGGCGACGTACGGAGTTGTGTACAAA 60
| | | | |
QY 61 TGCAGAAAACAAAACCTCTGACAGTAGTAGCTGTGTTAAAAAATTTGTGGAATCTGAAGAT 120
| | | | |
Db 61 GCCAGAAAACAAGTTGACGGGAGAGGTGGTGGCGCTTAAGAAAAATCCGCCIGGACACTGAG 120
| | | | |
QY 121 GATCCTGTGTTAAGAAAATAGCACTAAGAGAAAATACGTATGTTGAAGCAATTAACACAT 180
| | | | |
Db 121 ACTGAGGTGTGCCAGTACTGCCATCCGAGAGATCTCTCTGCTTAAGGAGCTTAACCAT 180
| | | | |
QY 181 CCARATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAGGAAAATGCATTTAGTTTTT 240
| | | | |
Db 181 CCTAATATTGTCAGCTGCTGGATGTCATTACACAGAAAATAAACTCTACCTGTTTTT 240
| | | | |
QY 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGA---AAGAAACCCAAATGGAGTTGCT 297
| | | | |
Db 241 GAATTTCTGCACCAAGATCTCAAGAAATTCATGGATGCTCTGCTCTCACTGGCATTCCT 300
| | | | |
QY 298 GATGGAGTGATCAAAAGCGTATTATATGGCAACACTTCAAGCTCTTAATTTCTGTATATA 357
| | | | |
Db 301 CTTCCCTCATCAAGAGCTATCTGTTCCAGCTGCTCCAGGCGCTAGCTTTCTGCAATTCT 360
| | | | |
QY 358 CATAACTGTATTACAGAGATATATAACCTGAAAATATTTCTAATAACTAAGCAAGGAATA 417
| | | | |
Db 361 CATCGGGTCTCCACCGAGACCTTAAACCTCAGAAATCTGCTTATTAAACACAGAGGGGCC 420
| | | | |
QY 418 ATCAAGATTTGTGACTTCGGGTTTGCACAAATCTCTGA---TTCCAGGAGATGCCCTACACC 474
| | | | |
Db 421 ATCAAGCTAGCAGACTTTGGACTAGCCAGAGCTTTTGGAGTCCCTGTTCTGACTTACACC 480
| | | | |
QY 475 GATTATGTAGCTACGAGATGGTACCGAGCTCCTGAACTTCTTGTGGGAGATACCTCAGTAT 534
| | | | |
Db 481 CATGAGGTGTCACCCCTGTGGTACCGAGCTCCTGAAAATCCTCCTGGGCTCGAAATATTAT 540
| | | | |
QY 535 GGTCTTTCAGTCGATATATATGGGCTATTGGTTGTGTTTTCAGAGAGCTCCTGACAGGCCAG 594
| | | | |
Db 541 TCCACAGCTGTGGACATCTGGAGCCTGGGCTGCATCTTTTGTGAGATGGTGAATCTCGCGG 600
| | | | |
QY 595 CCACTGTGGCCTGGAAAATCAGATGTGGACCAACTTTTATCTGATAATCAGAACACTAGGA 654
| | | | |
Db 601 GCCCTGTTCCTGGAGATTCTGAGATTGACCAGCTCTTCCGATCTTTCCGACTCTGGGG 660
| | | | |
QY 655 A 655
|
Db 661 A 661

RESULT 7
US-09-266-225D-13
; Sequence 13, Application US/09266225D
; Patent No. 6573364
; GENERAL INFORMATION:
; APPLICANT: Nandabalan, Krishan
; APPLICANT: Kingsmore, Stephen
; APPLICANT: Tchernev, Velizar
; TITLE OF INVENTION: Isolation and Characterization of Hermansky-Pudlak
; TITLE OF INVENTION: Syndrome (HPS) Protein Complexes and HPS Protein-
; TITLE OF INVENTION: Interacting Proteins
; FILE REFERENCE: 15966-523
; CURRENT APPLICATION NUMBER: US/09/266,225D
; CURRENT FILING DATE: 1999-03-10
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 1476
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-266-225D-13

Query Match 17.3%; Score 163.4; DB 4; Length 1476;
Best Local Similarity 55.1%; Pred. No. 1.5e-36;
Matches 364; Conservative 0; Mismatches 291; Indels 6; Gaps 2;


```

; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1632)
US-09-417-197-114

Query Match      17.3%; Score 163.4; DB 4; Length 1635;
Best Local Similarity 55.1%; Pred. No. 1.6e-36;
Matches 364; Conservative 0; Mismatches 291; Indels 6; Gaps 2;

QY 1 ATGGAAAGTATGAANAATTAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAA 60
Db 739 ATGGAGAACTTCCAAAGAGTGGAAAGATCGAGAGGCGACGTACGGAGTTGTGTACAA 798

QY 61 TGCAGAACAAACCTCTGGACAAAGTAGTAGCTGTTAAANAATTTGTGGATCTGRAGAT 120
Db 799 GCCAGAACAAAGTTGACGGGAGAGGTGGTGGCGCTTAAGANAATCCGCCCTGGACACTGAG 858

QY 121 GATCCTGTGTTAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAACAT 180
Db 859 ACTGAGGGTGTGCCAGTACTGCCATCCGAGAGATCTCTCTCTTAAGGAGCTTAACCAT 918

QY 181 CCAAACTCTGTGAACCTCATCGAGGTGTTCCAGGAGANAAGANAATGCAATTTAGTTTT 240
Db 919 CCTAATATGTCAGCTGCTGGATGTCATTCACACAGANAATAAATCTACTCTGTTTT 978

QY 241 GAATACCTGTGATCATACACTTTTAAATGAGCTGGA---AAGAAACCCAAATGGAGTTGCT 297
Db 979 GAATTTCTGCACCAAGATCTCAAGAAATTCATGGATGCTCTGCTCTCACTGGCATTCCT 1038

QY 298 GATGAGTGTATCAAAAGCGTATATATGGCAACACTTCAAGCTCTTAATTTCTGTCAATA 357
Db 1039 CTTCCTCTCATCAAGAGCTATCTGTTCCAGCTGCTCCAGGCGCTAGCTTTCTGCCATCT 1098

QY 358 CATAACTGTATTCACAGAGATATAAACCCTGANAATATTTCTAATAACTAAGCAAGGATA 417
Db 1099 CATCGGTCCTCCACCAGACCTTAACCTCAGAACTCTGCTTATTAACACAGAGGGGCC 1158

QY 418 ATCAAGATTTGTGACTTCGGGTTTGCACAAATTCCTGA---TTCCAGGAGATGCTACACC 474
Db 1159 ATCAAGCTAGCAGACTTTGGACTAGCCAGAGCTTTTGGAGTCCCTGTTCTGACTTACACC 1218

QY 475 GATTATGTAGCTACGAGATGGTACCGAGCTCCTGAACCTTCTTGGGAGATACCTAGTAT 534
Db 1219 CATGAGGTGTGACCTGTGGTACCGAGCTCCTGAATCTCCTCTGGCTCGAATATAT 1278

QY 535 GGTCTTCAGTCGATATATGGGCTATTGGTTGTGTTTTGCGAGAGCTCCTGACAGGCCAG 594
Db 1279 TCCACAGCTGGACATCTGGAGCCTGGGCTGCATCTTTGTGAGATGGTGACTGGCCGG 1338

QY 595 CCACTGTGGCTGGANAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGA 654
Db 1339 GCCCTGTTCCCTGGAGATTCTGAGATTGACCAGCTCTTCCGGATCTTTCGGACTCTGGG 1398

QY 655 A 655
Db 1399 A 1399

RESULT 10
US-09-220-132-3
; Sequence 3, Application US/09220132
; Patent No. 6506607
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT
; TITLE OF INVENTION: OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 07334-074001
; CURRENT APPLICATION NUMBER: US/09/220.132
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: US 60/079.303
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: US 60/068.821
; PRIOR FILING DATE: 1997-12-24
```

```

; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1050
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-220-132-3

Query Match      16.5%; Score 155.6; DB 4; Length 1050;
Best Local Similarity 54.4%; Pred. No. 2e-34;
Matches 361; Conservative 0; Mismatches 294; Indels 9; Gaps 2;

QY 1 ATGGAAAGTATGAANAATTAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAA 60
Db 127 ATGGAAGATTATACCAAAATAGAGAAATTTGGAGAAGGTACCTATGGAGTTGTATAG 186

QY 61 TGCAGAACAAACCTCTGGACAAAGTAGTAGCTGTTAAAAAATTTGTGGAATCTGAAGAT 120
Db 187 GGTAGACACAAACTACAGGTCAAGTGTAGCCATGAAAAAATCAGACTAGAAAGTGAA 246

QY 121 GATCCTGTGTTAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAACAT 180
Db 247 GAGGAAGGGGTTCCCTAGTACTGCAATTCGGGAAATTTCTCTATTAAAGGAACCTTCGTCAT 306

QY 181 CCAAACTCTGTGAACCTCATCGAGGTGTTCCAGGAGANAAGANAATGCAATTTAGTTTT 240
Db 307 CCAAAATATAGTCAGTCTTCAGGATGTGCTTATGCAGGATTCAGGTTATATCTCATCTT 366

QY 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGG-----AGTT 294
Db 367 GAGTTTCTTCCATGGATCTGAAGAAATACTTGGATTCTATCCCTCCTGGTCAGTACATG 426

QY 295 GCTGATGGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCAT 354
Db 427 GATTCCTCACTTGTAAAGAGTTATTTATACCAAAATCCTACAGGGGATTTGTTTTGTGTCAC 486

QY 355 ATACATAACTGTATTACAGAGATATAAACCCTGANAATATTTCTAATAACTAAGCAAGGA 414
Db 487 TCTAGAAGAGTTCTTCACAGAGACTTAANAACCTCAAAATCTCTTGATTGATGACAAAGGA 546

QY 415 ATAATCAAGATTTGTGACTTCGGGTTTGGACAAATTCGTG---ATTCCAGGAGATGCCTAC 471
Db 547 ACAATTAACCTGGCTGATTTTGGCCCTTGGCAGAGCTTTTGGAAATACCTATCAGAGTATAT 606

QY 472 ACCGATTATGTAGCTACGAGATGGTACCGAGCTCCTGAACTTCTTGTGGGAGATACCTCAG 531
Db 607 ACACATGAGGTAGTAACACTCTGTGTACAGATCTCCAGAAAGTATTGCTGGGGTCAGCTCGT 666

QY 532 TATGTTCTTCAGTCGATATATGGGCTATTGGTTGTGTTTTTGCAGAGCTCCTGACAGGC 591
Db 667 TACTCAACTCCAGTTGACATTTGGAGTATAGGCACCATATTTTGTGAACTAGCAACTAAG 726

QY 592 CAGCCACTGTGGCCTGGANAATCAGATGTGGACCAACTTTTATCTGATAATCAGAACACTA 651
Db 727 AAACCACCTTTTCCATGGGGATTCAGAAATTCAGAAATTCAGAACTCTTCAGGATTTTCAGAGCTTTG 786

QY 652 GGAA 655
Db 787 GGCA 790

RESULT 11
US-09-620-312D-313
; Sequence 313, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
```


Db	13	ATGCAGAAATACGAGAAACTGGAAAGATTTGGGAAGCGCACCTACGGAACCTGTGTTC AAG	72
Qy	61	TGCAGAAACAAAACCTCTGGACAAGTAGTACTGTTAAAAAATTTGTGGAATCTGAAGAT	120
Db	73	GCCAAAAACCGGAGACTCATGAGATCGTGGCTCTAAAAACGGGTGAGGCTGGATGACGAT	132
Qy	121	GATCCTGTGTTAAGAAAAATAGCACAATAAGAGAAAATACGTATGTTGAAGCAATTAAAAACAT	180
Db	133	GATGAGGGTGTCCGAGTTCGCGCCTCCGGGAGATCTGCCTACTCAAGGAGCTGAAGCAC	192
Qy	181	CCAAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAGGAAAAATGCATTTAGTTTTT	240
Db	193	AAGAACATCGTCAGGCTTCATGACGTCCTGCACAGCGCAAGAAAGCTGACTTTGGTTTTT	252
Qy	241	GAATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT	300
Db	253	GAATCTGTGACCAAGGACCTGAAGAAAGTATTTTGACAGTTGCAATGGTGACCTCGATCCT	312
Qy	301	GGAGTGATCAAAAGCGTATTATGGCAAAACACTTTC AAGCTCTTAAATTTCTGTCAATACAT	360
Db	313	GAGATTGTAAAGTCATTCTCTTCCAGCTACTAAAAGGGCTGGGATTCGTGCAZAGCCGC	372
Qy	361	AAC TGATTACAGAGATATTA A A A C C T G A A A A T A T T C T A A T A A C T A A G C A A G A A T A A T C	420
Db	373	AATGTGTACACAGGGACCTGAAGCCCCAGAACCTGTCTAATAAACAGGAAATGGGAGCTG	432
Qy	421	AAGATTGTGACTTCGGGTTTGCAAAA--TTCTGATCCAGGAGATGCCATACACCGAT	477
Db	433	AAATGGCTGATTTTGGCCGTGGCTCGAGCCTTTGGGATTCCTGCTCCGCTGTTACTCAGCT	492
Qy	478	TATGTAGCTACGAGATGGTACCGAGCTCCTGAACCTCTGTGGGAGATACCTAGTATGGT	537
Db	493	GAGGTGGTCACACTGTGGTACCGGCCACCGGATGTCTCTTTGGGGCCAAAGCTGTACTCC	552
Qy	538	TCTTCAGTCGATATATGGGCTATTGGTTGTGTTTTTGCAGAGCT--CCIGACAGGCCAG	594
Db	553	ACGTCCATCGACATGTGGTCAGCCGGCTGCATCTTTTCAGAGAGCTGGCCCAATGCTGGSCGG	612
Qy	595	CCACTGTGGCCCTGGAAAAATCAGATGTGGACCAACTTTTATCTGATAATCAGAAACACTAGGA	654
Db	613	CCTCTTTTCCCGGCAATGATGTCGATGACCCAGGATTTGAAAGAGGATCTTCCGACTGCTGGGG	672
Qy	655	A 655	
Db	673	A 673	

RESULT 13

US-08-464-517-37
; Sequence 37, Application US/08464517
; Patent No. 5869640
; GENERAL INFORMATION:
; APPLICANT: BEACH, David H.
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,517
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/963,308

QY 655 A 655
Db 673 A 673

RESULT 14

US-08-246-361A-37
; Sequence 37, Application US/08246361A
; Patent No. 5998582
; GENERAL INFORMATION:
; APPLICANT: BEACH, David H.
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/246,361A
; FILING DATE: 19-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/963,308
; FILING DATE: 16-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/888,178
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/701,514
; FILING DATE: 16-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Matthew P. Vincent
; REGISTRATION/DOCKET NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-004C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1089 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 13..888
US-08-246-361A-37

Query Match 14.2%; Score 134.6; DB 2; Length 1089;
Best Local Similarity 52.3%; Pred. NO. 1.8e-28;
Matches 346; Conservative 0; Mismatches 309; Indels 6; Gaps 2;

QY 1 ATGGAAGAATGATAAATAAGCTAGACTGGAGAGGGTCTTATGGGGTGTATTCAAA 60
Db 13 ATGCAGAAATACGAGAAATCTGGAAGAGATTGGGGAGAGCCCTACGGAACTGTGTTCAAG 72
QY 61 TGCAGAAACAAAACCTCTGAGCAAGTAGTAGCTGTGTAATAAATTTGTGGAATCTGAAGAT 120
Db 73 GCCAAAACCGGGAGACTCATGAGATCGTGGCTCTAAAACGGGTGAGGCTGGATGACGAT 132
QY 121 GATCCTGTTGTTAAGAAATAGCACTAAGAGAAATAGCTATGTTGAAGCAATTAAGCAT 180
Db 133 GATGAGGGTGTCCCGAGTTCGGCCCTCCGGGAGATCGCTACTCGCTACTCAAGGAGCTGAAGCAC 192

QY 181 CCAATCTTGTGAACCTCATCGAGGTGTTCCAGGAGAAAGGAAATGCAATTAGTTTTT 240
Db 193 AAGAAACATCGTCAGGCTTCATGACGTCCTGCACAGCAAGAGCTGACTTTGGTTTTT 252
QY 241 GAATACCTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGGAGTTGCTGAT 300
Db 253 GAATCTGTGACCAAGACCTGGAAGAGTATTTTGACAGTTGCAATGGTGACCTCGATCCT 312
QY 301 GGAGTGATCAAAAGCGTATTTATGGCAACACTTCAAGCTCTTAATTTCTGTGATATACAT 360
Db 313 GAGATTGTAAGTCAATTCCTCTTCCAGCTACTAAAGGGCTGGGATTCTGTGATAGCCGC 372
QY 361 AACTGTATTCAACAGAGATATAAAACCTGAAAATATCTTAATAACTAAGCAAGGAATAATC 420
Db 373 AATGTGCTACACAGGACCTGAAGCCCAAGAACTGCTAATAAACAGGAATGGGGAGCTG 432
QY 421 AAGATTGTGACTTCGGGTTTGCACAAA--TTCTGATTCAGGAGATGCTACACCCGAT 477
Db 433 AAATTGGCTGATTTTGGCTGCTCGAGCCCTTTGGGATTCCTGCTCGCTGTTACTCAGCT 492
QY 478 TATGTAGCTACGAGATGGTACCGAGCTCCTGAACCTTCTTGTGGGAGATACCTAGTATGGT 537
Db 493 GAGGTGGTCACTGTGTACCGCCCAAGGATGTCCTCTTTGGGGCCCAAGCTGACTACTCC 552
QY 538 TCTTCAGTCGATATATGGGCTATTTGGTGTGTTTGTGAGAGCT---CCTGACAGGCCAG 594
Db 553 ACGTCCATCGACATGTGTGTCAGCGGCTGCATCTTTGACAGAGCTGGCCCAATGCTGGCGG 612
QY 595 CCACTGTGGCCTGGAATAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGA 654
Db 613 CCTCTTTTCCCGGCAATGATGTCGATGACCAAGTGTGAAGAGGATCTTCCGACTGCTGGGG 672
QY 655 A 655
Db 673 A 673

RESULT 15

US-08-463-772-37
; Sequence 37, Application US/08463772
; Patent No. 6066501
; GENERAL INFORMATION:
; APPLICANT: BEACH, David H.
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,772
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/963,308
; FILING DATE: 16-OCT-1992
; APPLICATION NUMBER: US 07/888,178
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/701,514
; FILING DATE: 16-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Matthew P. Vincent
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-004C

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 1089 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 13..888
US-08-463-772-37

Query Match		14.2%;	Score 134.6;	DB 3;	Length 1089;
Best Local Similarity		52.3%;	Pred. No. 1.8e-28;		
Matches 346;		Conservative 0;	Mismatches 309;	Indels 6;	Gaps 2;
QY	1	ATGGAAAGTATGAAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGGTTCTATTCAAA	60		
DB	13	ATGCAGAAATACGAGAAACTGGAAGAAATTTGGGAGGGCACCTACGGAACCTGTGTTCAAG	72		
QY	61	TGCAGAAACAAAACCTCTGGACAGTAGTAGCTGTTAAAAAATTTGTGGAATCTGAAGAT	120		
DB	73	GCCAAAAACCCGGAGACTCATGAGATCGTGGCTCTAAACGGGTGAGGCTGSAATGACGAT	132		
QY	121	GATCCTGTGTTAAGAAATATGACATAAGAGAAATACGTATGTTGAAGCAATTAACACAT	180		
DB	133	GATGAGGGTGTGCCGAGTTCCGGCCCTCCGGGAGATCTGCCTACTCAAGGAGCTGAAGCAC	192		
QY	181	CCAAATCTTGTAACCTCATCGAGGTGTTTCAGGAGAAAGGAAATGCAATTTAGTTTTT	240		
DB	193	AAGAACATCGTCAGGCTTCATGACGTCTCTGCACACGGACAGAAAGCTGACTTTGGTTTTT	252		
QY	241	GAATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT	300		
DB	253	GAATTCCTGTGACCAGGACCTGAGAAAGTATTTTGACAGTTGCAATGGTGACCTCGATCCT	312		
QY	301	GGAGTGATCAAAAGCGTATTTATGGCAAAACACTTCAAGCTCTTAATTTCTGTATATACAT	360		
DB	313	GAGATTGTAAGTCAATTCCTCTTCCAGCTACTAAAGGGCTGGGATTCCTGCATAGCCGC	372		
QY	361	AACGTGATTACACAGAGATATAAAACCTGAAATATTTCTAATAACTAAGCAAGGAATAATC	420		
DB	373	AATGTGCTACACAGGGACCTGAGGCCCCAGAACCTGTCTAATAAACACAGGAATGGGAGCTG	432		
QY	421	AAGATTGTGACTTCGGGTTTGCAAAA---TTCTGATTCACAGGAGATGCCCTACACCGAT	477		
DB	433	AAATTGGCTGATTTTGGCCTGGCTCGAGCCTTTGGGATTCGCCCTCCGCTGTTACTCAGCT	492		
QY	478	TATGTAGCTACGAGATGGTACCGAGCTCCTGAACCTTCTTGTGGGAGATACTCAGTATGGT	537		
DB	493	GAGGTGGTCACACTGTGGTACCGCCACCGGATGTCTCTTTGGGGCCCAAGCTGTACTCC	552		
QY	538	TCTTCAGTCGATATATGGGCTATTGGTTGTGTTTTTGCAGAGCT---CCTGACAGGCCAG	594		
DB	553	ACGTCCATCGACATGTGGTCAGCCGGCTGCATCTTTGACAGAGCTGGCCAATGCTGGGCGG	612		
QY	595	CCACTGTGGCTGGAATAATCAGATGTGGACCAACTTTATCTGATAATCAGAAACACTAGGA	654		
DB	613	CCTCTTTTCCCGCAATGATGTGATGACCAAGTTGAAGAGGATCTTCCGACTGCTGGGG	672		
QY	655	A 655			
DB	673	A 673			

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2003, 16:06:46 ; Search time 2379.15 Seconds
(without alignments)
9653.738 Million cell updates/sec

Title: US-09-671-050-11
Perfect score: 945
Sequence: 1 atggaaaagtatgaaaaatt.....aggtacttcgcgtcaaaagt 945

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	480.6	50.9	695	14 CB169554	CB169554 RUC603000
2	405	42.9	670	14 BY733578	BY733578 BY733578
3	401.8	42.5	1691	11 AK016781	AK016781 Mus muscu
4	390.6	41.3	825	14 CB315094	CB315094 AGENCOURT

C	5	361	38.2	536	13	BU686325	BU686325 UI-CF-DUI
	6	360.2	38.1	689	14	BY752739	BY752739 BY752739
	7	333.6	35.3	500	9	AI385966	AI385966 ml34h09.Y
	8	319.4	33.8	824	13	BU221831	BU221831 603750354
	9	316.8	33.5	938	9	AA061797	AA061797 ml34h09.Y
	10	306.8	32.5	598	9	AV986182	AV986182 AV986182
	11	289.2	30.6	597	12	BP016186	BP016186 BP016186
	12	269.8	28.6	565	14	CB400506	CB400506 OSTF177F1
	13	258.6	27.4	731	9	AW106692	AW106692 um32d03.Y
C	14	255	27.0	645	14	CA334835	CA334835 NISC lt01
	15	255	27.0	718	9	AA286088	AA286088 vc33b05.Y
	16	250.6	26.5	2572	11	AK030598	AK030598 Mus muscu
	17	249	26.3	508	12	BI511582	BI511582 BS160006A
	18	247.8	26.2	512	13	BX304025	BX304025 BX304025
	19	239	25.3	485	9	AV960213	AV960213 AV960213
	20	231.2	24.5	812	13	BU377342	BU377342 603812152
C	21	227.2	24.0	1064	13	BU120363	BU120363 603142234
	22	226	23.9	774	13	BW110560	BW110560 BW110560
	23	224.2	23.7	2005	11	AK090045	AK090045 Mus muscu
	24	224.2	23.7	2769	11	AK045036	AK045036 Mus muscu
	25	223.8	23.7	688	13	BW241505	BW241505 BW241505
	26	222.6	23.6	1893	11	AK050990	AK050990 Mus muscu
	27	221	23.4	2388	11	AK046394	AK046394 Mus muscu
	28	221	23.4	2665	11	AK045356	AK045356 Mus muscu
	29	218.6	23.1	750	14	CB955302	CB955302 AGENCOURT
	30	216.2	22.9	432	14	CB757579	CB757579 AMGNNUC:N
	31	213.4	22.6	570	12	BJ000534	BJ000534 BJ000534
	32	211.6	22.4	2013	11	AK078804	AK078804 Mus muscu
	33	210.8	22.3	627	14	BY716459	BY716459 BY716459
	34	209.8	22.2	579	9	AW233105	AW233105 fj28b10.Y
	35	209.4	22.2	380	13	BY075395	BY075395 BY075395
C	36	208.4	22.1	2604	11	AK052380	AK052380 Mus muscu
	37	203.4	21.2	295	9	AA626859	AA626859 zu89f10.s
	38	197.2	20.9	471	14	CB732963	CB732963 AMGNNUC:N
	39	196.9	20.8	851	13	BU463954	BU463954 603369666
	40	192.4	20.4	1535	11	BC010966	BC010966 Homo sapi
	41	188.6	20.0	659	13	BW280090	BW280090 BW280090
	42	187	19.8	673	13	BW036816	BW036816 BW036816
	43	186.6	19.7	727	9	AI508835	AI508835 vc33b05.Y
	44	186.2	19.7	666	14	BY733092	BY733092 BY733092
	45	182	19.3	320	13	BY121897	BY121897 BY121897

ALIGNMENTS

RESULT 1
CB169554/c
LOCUS CB169554 695 bp mRNA linear EST 30-JAN-2003
DEFINITION RUC603000752.R1 CSEQFXN20 testes Bos taurus cDNA, mRNA sequence.
ACCESSION CB169554
VERSION CB169554.1 GI:28155681
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus

REFERENCE
AUTHORS Adelson,D.L. and Gill,C.A.
TITLE Bovine ESTs (Adelson and Gill)
JOURNAL Unpublished
COMMENT Contact: David L. Adelson
Animal Breeding and Genetics
Texas A&M University
Animal Science Dept., TAMU-2471, College Station, TX 77843-2471,
USA
Tel: 9798452616
Fax: 9798456970
Email: david.adelson@tamu.edu.
Location/Qualifiers
1. .695
/organism="Bos taurus"

FEATURES
source

```

/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="testes"
/clone_lib="CSEQFXN20 testes"
/note="Organ: testes; Vector: pBluescript SK+; Site_1:
NotI; Site_2: EcoRI; sequence 5' of the insert
(5'-NNN...NNNinsert)
GCGAATTCGAGCTCCACCGCGTGGCGGGCCGGCTCGAG. Sequence 3' of
the inserts (AAGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG.
normalized Rd 1 library, sequenced 3' with M13R primer."
BASE COUNT      169 a   166 c   133 g   227 t
ORIGIN
```

Query Match	50.9%	Score 480.6;	DB 14;	Length 695;
Best Local Similarity	92.8%	Pred. No. 5.9e-96;		
Matches 504;	Conservative 0;	Mismatches 39;	Indels 0;	Gaps 0;
QY	1	ATGGAAAGTATGAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGTGTATTCAA	60	
DB	543	ATGGAGAAGTATGAAAATTAGCTAAGATTGGGAGGGTCTTATGGGTGTATTCAA	484	
QY	61	TGCAGAAACAAACCTCTGGACAAGTAGTAGTGTAAATAATTTGTGAATCTGAAGAT	120	
DB	483	TGCAGAAACAAACCTCTGGACAAGTGGTAGCTATTAAAAAATTTGTGAATCCGAAGAT	424	
QY	121	GATCCTGTTGTTAAGAAAATAGCACTAAGACAAATACGTATGTTGAAGCAATTAAACAT	180	
DB	423	GATCCTGTCGTTAAGAAAATAGCTCTAAGAGAAAATACGCATGTTAAGCAATTAAACAT	364	
QY	181	CCAAATCTTGTGAACCTCATCGAGGTGTTCCAGAGAAAAGGAAATGCATTAGTTTT	240	
DB	363	CCAAATCTTGTGAACCTCATCGAGGTGTTCCAGAGAAAAGGAAATGCATTAGTTTT	304	
QY	241	GAATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT	300	
DB	303	GASTACTGTGATCACACACTTTTAAATGAGCTGGAAAGAAACCCAAACGGAGTTGCTGAT	244	
QY	301	GGAGTGATCAAAAGCGTATTATATGGCAACACTTCAAGCTCTTAATTTCTGTCAATACAT	360	
DB	243	GGAGTGATCAAAAGTGTGTATGGCAACACTCCAAGCTCTCAACTCTGTCAATAACAC	184	
QY	361	AACTGTATTCACAGAGATATAAAACCTGAAATATTCTAATAACTAAGCAAGGAATAATC	420	
DB	183	AATTGTATTCACAGAGATGTAAAACCTGAAATATTCTAATAACTAAGCAAGGAATAATC	124	
QY	421	AAGATTGTGACTTCGGGTTGCACAAATTCGATCCAGGAGATGCTACACCGATTAT	480	
DB	123	AAGATCTGTGACTTTGGGTTGCACGAATTCGTATCCAGGAGATGCTACACTGACTAC	64	
QY	481	GTAGCTACGAGATGGTACCGAGCTCCCTGAATCTTTGTGGGAGATACAGTATGGTTCT	540	
DB	63	GTCGCTACAAGATGGTACCGAGCTCCCGACCTTTCTGTGGGGATACCCAGTATGGTTCT	4	
QY	541	TCA 543		
DB	3	TCA 1		

```

RESULT 2
BY733578
LOCUS      BY733578      670 bp      mRNA      linear      EST 17-DEC-2002
DEFINITION BY733578 RIKEN full-length enriched, 16 days neonate male
            diencephalon Mus musculus cDNA clone G630052E12 5', mRNA sequence.
ACCESSION  BY733578
VERSION     BY733578.1  GI:27146705
KEYWORDS
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 670)
AUTHORS     Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
            Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I., Kiyosawa,H.
```

Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A., Schonbach,C., Gojohori,T., Baldarelli,R., Hill,D.P., Bult,C., Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H., Batalov,S., Beisel,K.W., Blake,S.A., Bradt,D., Brusic,V., Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A., Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T., Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S., Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A., Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R., Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T., Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S., Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M., Sandelin,A., Schneider,C., Sempile,C.A., Setou,M., Shimada,K., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M., Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y., Wells,C., Wilming,L.G., Wyrshaw-Boris,A., Yanagisawa,M., Yang,I., Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P., Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K., Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S., Rogers,J., Birney,E. and Hayashizaki,Y.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

22354683

12466851

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gsc.riken.go.jp,

URL:http://genome.gsc.riken.go.jp/

Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M., Takeda,Y., Waki,K., Wataniki,A., Muramatsu,M. and Hayashizaki,Y. Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

Location/Qualifiers

1. .670

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="G630052E12"

/sex="male"

/tissue_type="diencephalon"

FEATURES

source

```
/dev stage="16 days neonate"
/clone_lib="RIKEN full-length enriched, 16 days neonate
male diencephalon"
BASE COUNT 205 a 140 c 169 g 155 t 1 others
ORIGIN
Query Match 42.9%; Score 405; DB 14; Length 670;
Best Local Similarity 85.6%; Pred. No. 3.2e-79;
Matches 450; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

Qy 1 ATGGAAGTATGAAATAGCTAAGACTGGAGAGGGTCTTATGGGGTGTATTCAA 60
Dd ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
145 ATGGAAGTATGAAAGCTAGCTAAGATCGGAGAGGGTCTTATGGCGTGTATTCAAG 204

Qy 61 TGCAGAAACAAACCTCTGGACAAGTAGTAGCTGTAAATAATTTGTGGAATCTGAAGAT 120
Dd ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
205 TGCAGAAACAAATCTTCTGGACAAGTAGTAGCGATCAAAAATTCGTGGAATCTGAAGAT 264

Qy 121 GATCCTGTTGTTAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAT 180
Dd ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
265 GATCGTGTGTTAGGAAATAGCCCTGCGGAAATCCGTATGCTGAAGCAGTTGAAACAC 324

Qy 181 CCAATCTTGTGAACCTCATCGAGGTGTTTCAAGGAGAAAGGAAATGCAATTTAGTTTTT 240
Dd ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
325 CCAACCTCGTGAACCTCATCGAGGTGTTTCAAGGAGAAAGGAAATGCAATTTAGTTTTT 384

Qy 241 GAATAGTGTGATCATACACTTTTAAATGAGCTGGAGAAACCCAAATGGAGTTGCTGAT 300
Dd ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
385 GAGTACTGTGATCACACACTGTTAAACGAGCTGGAGAGAAACCCAAACGAGTTTCTGAT 444

Qy 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCAATACAT 360
Dd ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
445 GGAGTGATTAAGAGTGTGATGTCGCAACCTTCAAGCCCTTAACCTTCTGTCAACGAC 504

Qy 361 AACTGTATTACAGAGATATAAAACCTGAAATATTCTAATACTAAGCAAGGAATATC 420
Dd ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
505 AATTGTATTATTCGGGATGTAAACCTTGAAACATCTTAATACCAAGCAAGGGATGATA 564

Qy 421 AAGATTGTGACTTCGGGTTTGACAAATTTCTGATCCAGGAGATGCTACACCGATTAT 480
Dd ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
565 AAGATTGTGACTTTGGATTGACGGAATTTCTAATCCAGGAGACGCTACACAGACTAT 624

Qy 481 GTAGTACGAGATGGTACCGAGCTCTGAACTTCTGTGGGAGATA 526
Dd ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
625 GTTCCACCAAGGTGGTACCGAGCCNCGAACTTCTGTGGGAGACA 670

RESULT 3
AK016781
LOCUS
DEFINITION
Mus musculus adult male testis cDNA, RIKEN full-length enriched
library, clone:4933411017 product:cyclin-dependent kinase-like 1
(CDC2-related kinase), full insert sequence.
AK016781 1691 bp mRNA linear HTC 05-DEC-2002
AK016781.1 GI:12855701
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374

Qy 1 ATGGAAGTATGAAATAGCTAAGACTGGAGAGGGTCTTATGGGGTGTATTCAA 60
Dd ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
145 ATGGAAGTATGAAAGCTAGCTAAGATCGGAGAGGGTCTTATGGCGTGTATTCAAG 204

Qy 61 TGCAGAAACAAACCTCTGGACAAGTAGTAGCTGTAAATAATTTGTGGAATCTGAAGAT 120
Dd ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
205 TGCAGAAACAAATCTTCTGGACAAGTAGTAGCGATCAAAAATTCGTGGAATCTGAAGAT 264

Qy 121 GATCCTGTTGTTAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAT 180
Dd ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
265 GATCGTGTGTTAGGAAATAGCCCTGCGGAAATCCGTATGCTGAAGCAGTTGAAACAC 324

Qy 181 CCAATCTTGTGAACCTCATCGAGGTGTTTCAAGGAGAAAGGAAATGCAATTTAGTTTTT 240
Dd ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
325 CCAACCTCGTGAACCTCATCGAGGTGTTTCAAGGAGAAAGGAAATGCAATTTAGTTTTT 384

Qy 241 GAATAGTGTGATCATACACTTTTAAATGAGCTGGAGAAACCCAAATGGAGTTGCTGAT 300
Dd ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
385 GAGTACTGTGATCACACACTGTTAAACGAGCTGGAGAGAAACCCAAACGAGTTTCTGAT 444

Qy 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCAATACAT 360
Dd ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
445 GGAGTGATTAAGAGTGTGATGTCGCAACCTTCAAGCCCTTAACCTTCTGTCAACGAC 504

Qy 361 AACTGTATTACAGAGATATAAAACCTGAAATATTCTAATACTAAGCAAGGAATATC 420
Dd ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
505 AATTGTATTATTCGGGATGTAAACCTTGAAACATCTTAATACCAAGCAAGGGATGATA 564

Qy 421 AAGATTGTGACTTCGGGTTTGACAAATTTCTGATCCAGGAGATGCTACACCGATTAT 480
Dd ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
565 AAGATTGTGACTTTGGATTGACGGAATTTCTAATCCAGGAGACGCTACACAGACTAT 624

Qy 481 GTAGTACGAGATGGTACCGAGCTCTGAACTTCTGTGGGAGATA 526
Dd ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
625 GTTCCACCAAGGTGGTACCGAGCCNCGAACTTCTGTGGGAGACA 670

REFERENCE
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
```

```
11042159
3
PUBMED
REFERENCE
AUTHORS
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuina, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4
PUBMED
REFERENCE
AUTHORS
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
Saito, T., Okazaki, Y., Gojobori, T., Boro, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, M., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G.,
Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,
Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seva, T., Shibata, Y., Storch, K. F., Suzuki, H.,
Toyo-oka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohtsuki, S.
and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6921), 685-690 (2001)
21085660
11217851
5
PUBMED
REFERENCE
AUTHORS
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 1691)
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,
Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,
Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,
Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K.,
Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
Please visit our web site (http://genome.gsc.riken.go.jp/) for
further details.
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
[5' GAGAGAGAGAGGATCCAGAGCTCTTTTCTTTTCTTNN 3'], cDNA was
```


(where B = A, C, or G and N = A, C, G, or T). Average insert size 1.23 kb (range 0.5-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

BASE COUNT 234 a 188 c 200 g 202 t ; others
ORIGIN

Query Match 41.3%; Score 390.6; DB 14; Length 825;
Best Local Similarity 83.3%; Pred. No. 5.2e-76;
Matches 455; Conservative 0; Mismatches 90; Indels 1; Gaps 1;

QY 361 AACTGTATTCACAGAGATATAAAACCTGAAATATTTCTAATAACTAAGCAGGAATAATC 420
|||
Db 96 AGCTGTATCCATCGGACGCTAAACCCGAAACATCTCTAATAGCCAAAGCAGGTGATA 155
|||

QY 421 AAGATTGTGACTTCGGGTTTGGACAAATTTCTGATTCAGGAGATGCCTACACCGATTAT 480
|||
Db 156 AAGATTGTGACTTCGGATTTCAGCAATTTCTAATTCAGGAGATGCCTACACAGACTAT 215
|||

QY 481 GTAGCTACGAGATGGTACCGAGCTCCTGAACTTTCTGTGGGAGATACTCAGTATGGTTCT 540
|||
Db 216 GTTGCCACCCAGGTGTTACCGAGCCCTGAACTTTCTTGGGAGACACGAAGTACGGCTCC 275
|||

QY 541 TCAGTCGATATATGGGCTATTTGGTGTGTTTTTTCAGAGCTCCTGACAGGCCAGCCACTG 600
|||
Db 276 TCTGGACATATGGGCTGTGCGCTGTGTTTTTGCAGAGCTCCTGACCGGTCAGCCACTG 335
|||

QY 601 TGGCCCTGGAATAATCAGATGTGACCAACTTTATCTGATTAATCAGAACACATAGGAAATTA 660
|||
Db 336 TGGCCAGGAAATCAGATGTGACCAACTTTATCTGATCATCAGGACATTTGGGAAAGCTG 395
|||

QY 661 ATCCCAAGACATCAATCAATCTTTAAAGTAACGGGTTTTTCCATGGCATCAGTATACCT 720
|||
Db 396 ATCCCAAGGCATCAGTCTATCTTTAAAGTAACCCAGTTTTCGGTGGCATCAGCATACCC 455
|||

QY 721 GAGCCAGAAGACATGGAACTCTTGAGGAAAGTTCTCAGATGTTTCATCCCTGGCTCTG 780
|||
Db 456 GAACCTGAGGACATGGAGACTCTTGAGGAAATACTCTCAATGTTTCAGCCCTATGGCTTTA 515
|||

QY 781 AACTTCATG-AAGGGGTGCTGAAGATGAATCCAGATGACAGATTAACTGTTCCCAACT 839
|||
Db 516 AGTTTCATGAAGGGGTGCTGAGATGGATCCTGTGATGAGGCTGACCTGTGCCAGCT 575
|||

QY 840 CTTGGAGAGCTCCTACTTTTGATTCTTTTCAAGAGGCCCAATTAAGAAAGCAGGTAA 899
|||
Db 576 GTTGGACACTACCTACTTCAGTCTTTTCAAGAGGCCCAATTAAGAAAGCAGGTAA 635
|||

QY 900 TGAAGG 905
|||
Db 636 TGAGGG 641
|||

RESULT 5
BU686325/c
LOCUS
DEFINITION
UI-CF-DUI-adn-i-10-0-UI.s1 UI-CF-DUI Homo sapiens cDNA clone
536 bp mRNA linear EST 07-OCT-2002
UI-CF-DUI-adn-i-10-0-UI 3', mRNA sequence.

ACCESSION
BU686325
VERSION
BU686325.1 GI:23541120
KEYWORDS
EST.

SOURCE
Homo sapiens (human)
Homo sapiens

ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 (bases 1 to 536)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
AUTHORS
Normalization and subtraction: two approaches to facilitate gene
TITLE
discovery

JOURNAL
Genome Res. 6 (9), 791-806 (1996)
MEDLINE
97044477
PUBMED
8889548
COMMENT
Contact: McCray, PB

McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com) or from Open Biosystems
(www.openbiosystems.com).
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
source
1. .536
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-DUI-adn-i-10-0-UI"
/tissue_type="Primary Lung Epithelial Cells"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-CF-DUI"
/note="Organ: Lung; Vector: p77T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-CF-DUI is a normalized cDNA library containing the
following tissue(s): Primary Lung Epithelial Cells The
library was constructed according to Bonaldo, Lennon and
Soares, Genome Research, 6:791-806, 1996. First strand
cDNA synthesis was primed with an oligo-dT primer
containing a Not I site. Double stranded cDNA was ligated
to an EcoR I adaptor, digested with Not I, and cloned
directionally into p77T3-Pac vector. The oligonucleotide
used to prime the synthesis of first-strand cDNA contains
a library tag sequence that is located between the Not I
site and the (dT)18 tail. The sequence tag for this
library is GGCTGTAGGC.
TAG LIB=UI-CF-DUI
TAG_TISSUE=Lung Epithelial Cells Tissue nos 359-368
TAG_SEQ=GGCTGTAGGC"

BASE COUNT 148 a 104 c 97 g 187 t
ORIGIN

Query Match 38.2%; Score 361; DB 13; Length 536;
Best Local Similarity 100.0%; Pred. No. 1.7e-69;
Matches 361; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 585 GACAGGCCAGCCACTGTGGCTGGAAATCAGATGTGGACCACTTTATCTGATAATCAG 644
|||
Db 536 GACAGGCCAGCCACTGTGGCTGGAAATCAGATGTGGACCACTTTATCTGATAATCAG 477
|||

QY 645 AACACTAGGAAATTAATCCCAAGACATCAATCTTTAAAGTAACGGGTTTTTCCA 704
|||
Db 476 AACACTAGGAAATTAATCCCAAGACATCAATCTTTAAAGTAACGGGTTTTTCCA 417
|||

QY 705 TGGCATCAGTATACCTGAGCCAGACATGGAACTCTTGAGGAAAGTTCTCAGATGT 764
|||
Db 416 TGGCATCAGTATACCTGAGCCAGACATGGAACTCTTGAGGAAAGTTCTCAGATGT 357
|||

QY 765 TCATCCTGTGGCTCTGAACTTCATGAAGGGGTGTCTGAAGATGAATCCAGATGACAGATT 824
|||
Db 356 TCATCCTGTGGCTCTGAACTTCATGAAGGGGTGTCTGAAGATGAATCCAGATGACAGATT 297
|||

QY 825 AACCTGTTCCCAACTCCTGGAGAGCTCTACTTTGATTCTTTCAAGAGGCCCAATTAA 884
|||
Db 296 AACCTGTTCCCAACTCCTGGAGAGCTCTACTTTGATTCTTTCAAGAGGCCCAATTAA 237
|||

QY 885 AAGAAAAGCAGTAATGAAGGAAGAAACAGAGACGCCAACAGGTACTTCCGCTCAAAAG 944
|||
Db 236 AAGAAAAGCAGTAATGAAGGAAGAAACAGAGACGCCAACAGGTACTTCCGCTCAAAAG 177
|||

QY 945 T 945

Db 176 T 176

RESULT 6

BY752739

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

BY752739 689 bp mRNA linear EST 17-DEC-2002
BY752739 RIKEN full-length enriched, adult inner ear Mus musculus
CDNA clone F930014I06 5', mRNA sequence.

BY752739 Gi:27183802

EST.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 689)

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H.,
Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C.,
Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C., Hume, D. A.,
Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batalov, S.,
Beisel, K. W., Blake, J. A., Bradt, D., Brusic, V., Chothia, C., Corbani,
L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest,
A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,
Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, J. J.,
Jarvis, E. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R. M.,
King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons,
P. A., Maglott, D. R., Maltais, L., Marchionni, J., McKenzie, L., Miki,
H., Nagashima, T., Numata, K., Okido, T., Pavan, M. J., Pertea, S.,
Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. J., Qi, D.,
Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring,
B. Z., Ringwald, M., Sandelin, A., Schneider, C., Sempile, C. A., Setou,
M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale,
R. D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y.,
Watanabe, Y., Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa,
M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A.,
Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura,
M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii,
Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata,
K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander,
E. S., Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

22354683

12466851

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center(GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gs.riken.go.jp,

URL: http://genome.gsc.riken.go.jp/

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda,
S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F., Imotani, K.,
Ishii, Y., Itoh, M., Kagawa, I., Kawai, C., Kojima, Y., Kondo, S., Konno,
H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K.,
Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N., Sano, H.,
Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y.,
Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct
Submission.

Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Kirk W. Beisel (Boys Town National
Research Hospital 555 North 30th Street Omaha, NE 68131 USA) whose
assistance we gratefully acknowledge.
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.

Location/Qualifiers

1. 689
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="F930014I06"
/tissue_type="inner ear"
/dev_stage="adult"
/clone_lib="RIKEN full-length enriched, adult inner ear"
213 a 133 c 171 g 168 t 4 others

BASE COUNT
CRIGIN

Query Match 38.1%; Score 360.2; DB 14; Length 689;
Best Local Similarity 82.7%; Pred. No. 2.7e-69;
Matches 421; Conservative 0; Mismatches 87; Indels 1; Gaps 1;

QY 1 ATGGAAAAGTATGAAAAATTAGCTAAGACTGGAGAAAGGGTCTTATGGGGTTGTATTCAA 60
|||||
Db 169 ATGGAAAAGTATGAAAAGCTAGCTAAGATCGGAGAAAGGGTCTTATGGCGTTGTATCAAG 228
|||||
QY 61 TGCAGAAACAAAACCTCTGGACAGTAGTAGCTGTGTTAAAAAATTTGTGGAATCTGAAGAT 120
|||||
Db 229 TGCAGAAACAAAATCTTCTGGACAAGTAGTAGCGATCAAAAAAATTCGTGGAATCTGAAGAT 288
|||||
QY 121 GATCCTGTTGTTAAGAAAATAGCACTAAGAGAAATACGTTATGTTGAAGCAATTTAAACAT 180
|||||
Db 289 GATCGTGTCTGTAGGAAAATAGCCCTCGGGAAATCCGTTATGTTGAAGCAGTTGAACAC 348
|||||
QY 181 CCAATCTGTGAAACCTCATCGAGGTGTTTCAGGAGAAAAGGAAAAATGCATTTAGTTT 240
|||||
Db 349 CCAACCTCGTGAACCTCATTTAGGTGTTTCAGAGAGAAAGAGGAAGATGCATCTAGTTT 408
|||||
QY 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGGAGTTGCTGAT 300
|||||
Db 409 GAGTACTGTGATCACACGCTGTAAACGAGCTGGAGAGAAACCCAAACGGAGTTTCTGAT 468
|||||
QY 301 GGAGTGATCAAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGTCATATACAT 360
|||||
Db 469 GGAGTGATTAAGAAGTGTCTATGSCAAACCCCTTCAAGCCCTTAACTTCTGTCACAGCAC 528
|||||
QY 361 AACTGTATTCACAGAGATATAAAACCTGAAAATATTTCTAATAACTAAGCAAGGAATAATC 420
|||||
Db 529 AATTGTATTTCATCGGGATGTAAACCTGAAAACATCTCTAATAACCAAGCAGGGATGANT 588
|||||
QY 421 AAGATTTGTGACTTCGGGTTTGCAAAAATCTCTGATTCAGGAGATGCCTACACCGATTAT 480
|||||
Db 589 AAGATTTGTGACTTTGGATTGGACGAATTTCTAATTNCA-GAGACGCTTACACANGACTA 647
|||||
QY 481 GTAGCTACGAGATGTTACCGAGCTCCTGA 509
|||||
Db 648 TGTGCCANCCAGTGTACCGAGCCCCCGA 676

RESULT 7

AI385966

LOCUS

DEFINITION

AI385966 500 bp mRNA linear EST 27-JAN-1999
ml34h09.y1 Stratagene mouse testis (#937308) Mus musculus CDNA
clone IMAGE:513953 5' similar to gb:X66358 SERINE/THREONINE-PROTEIN

KINASE KKIALRE (HUMAN);, mRNA sequence.

AI385966
AI385966.1 GI:4199429
EST.

Mus musculus (house mouse)
Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 500)

Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wyllie,T.,
Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter
,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
Waterston,R. and Wilson,R.

The WashU-NCI Mouse EST Project 1999
Unpublished

Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:307801

This read is a RESEQUENCE of a previously sequenced mouse clone
correct orientation)

Seq primer: -40RP from Gibco

High quality sequence stop: 339.

Location/Qualifiers

1..500
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CD-1"
/db_xref="taxon:10090"
/clone="IMAGE:513953"
/sex="males"
/tissue_type="testis"
/dev_stage="10-12 week old"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="Stratagene mouse testis (#937308)"
/note="Organ: testis; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. Average insert size: 1.0 kb; Uni-Zap XR Vector;
-5' adaptor sequence: 5' GAATTCGGCAGAG 3' -3' adaptor
sequence: 5' CTCGATTTTCTTTTCTTTT 3'"

BASE COUNT 145 a 114 c 121 g 120 t

ORIGIN

Query Match 35.3%; Score 333.6; DB 9; Length 500;
Best Local Similarity 82.3%; Pred. No. 2e-63;
Matches 408; Conservative 0; Mismatches 84; Indels 4; Gaps 2;

QY 135 GAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAAACATCCAAATCTTGTGAA 194
|||||
Db 2 GAAATAGCCCTGCGGGAATCCGTATGCTGAAG---TTGAACACCCCAACCTCGTGAA 58
|||||

QY 195 CCTCATCGAGGTGTTCCAGGAGAAAAGGAAATGCATTAGTTTTTGAATACGTGTATCA 254
|||||

Db 59 CCTCATCGAGGTGTTCCAGAGAAAAGAGAAAGATGCATCTAGTTTTTGAGTACTGTGATCA 118
|||||

QY 255 TACACTTTTAAATGAGCTGGAAGAAACCCCAATGGAGTTGCTGATGGAGTGATCAAAAAG 314
|||||

Db 119 CACACTGTTAACGAGCTGGAGAGAAACCCCAACGGAGTTCTGTGATGGAGTGATTAAG 178
|||||

QY 315 CGTATTATGGCAACACTTCAAGCTCTTAATTTCTGCATATACATAAATCTGATTCACAG 374
|||||

Db 179 TGTGCTAIGGCAAAACCTTCAAGCCCTTAACCTTCTGTCAAGCACAATTTGATTCATCG 238
|||||

QY 375 AGATATAAAACCTGAAATAATTTCTAATAACTAAGCAAGGAATAATCAAGATTTCTGACTT 434
|||||

Db 239 GGATGTAAACCTGAAACATCCTTAATAACCAAGCAAGGGATGATAAAGATTTGTGACTT 298
|||||

QY 435 CGGGTTTGCAAAATTTCTGATTCAGGAGATGCCTACACCGATTATGTAGTACGATG 494
|||||

Db 299 TGGATTTGCACGAATTTCTAATTCAGGAGACGCCTACACAGACTATGTTGCCACCGTG 358
|||||

QY 495 GTACCGAGCTCTGAACTTTCTTGTGGGAGATACTCAGTATGGTTTCTCAGTCGATATATG 554
|||||

Db 359 GTACCGAGCCCTCGAACTTCTCGTGGGAGACACGAAGTAGCGTTCTCTGTAGACGTGTG 418
|||||

QY 555 GGCTATTGTTGTGTTTTTGCAGAGCTCTGACAGGCCAGCCACTGTGGCCTGGAATAATC 614
|||||

Db 419 GGCCGTCGGCTGTGTTTATGCAGAGCTTCTGACGTGTGCGCACTCTGGCCCGG-AAATA 477
|||||

QY 615 AGATGTGGACCAACTT 630
|||||

Db 478 CGACGTGGACCACTT 493
|||||

RESULT 8
BU221831

LOCUS
DEFINITION BU221831 824 bp mRNA linear EST 25-NOV-2002
603750354F1 CSEQCHN04 Gallus gallus cdna clone ChEST661113 5', mRNA
sequence.

ACCESSION BU221831
VERSION BU221831.1 GI:25410266

KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus

REFERENCE 1 (bases 1 to 824)
Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.

TITLE A Comprehensive Collection of Chicken cDNAs
JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)
MEDLINE 22335534
PUBMED 12445392

COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology (UMIST)
; PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers
1..824
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="ChEST661113"
/tissue_type="whole embryo"
/dev_stage="20-21"
/lab_host="DH108"
/clone_lib="CSEQCHN04"
/note="Organ: whole embryo; Vector: pBluescript II KS(+);
Site_1: EcoRI; Site_2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunted, ligated to NotI adapters, digested with EcoRI
, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."

BASE COUNT 192 a 227 c 239 g 166 t

ORIGIN

LOCUS AV986182 598 bp mRNA linear EST 14-MAR-2002
DEFINITION AV986182 Nori Satoh unpublished cDNA library, larva Ciona
intestinalis cDNA clone cilv41b18 5', mRNA sequence.
ACCESSION AV986182
VERSION AV986182.1 GI:19475050
KEYWORDS EST.
SOURCE Ciona intestinalis
ORGANISM Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Cionidae; Ciona.
REFERENCE 1 (bases 1 to 598)
AUTHORS Satoh,N., Satou,Y., Kohara,Y. and Shin-i,T.
TITLE Expressed genes in Ciona intestinalis
JOURNAL Unpublished
COMMENT Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoheascidian.zool.kyoto-u.ac.jp.
FEATURES
source
1..598
/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="cilv41b18"
/tissue_type="whole animal"
/dev_stage="larva"
/clone_lib="Nori Satoh unpublished cDNA library, larva"
BASE COUNT 179 a 105 c 149 g 165 t
ORIGIN
Query Match: 32.5%; Score 306.8; DB 9; Length 598;
Best Local Similarity 70.7%; Pred. No. 1.3e-57;
Matches 423; Conservative 0; Mismatches 172; Indels 3; Gaps 1;
QY 11 ATGAAATAATTAGCTAAGACTGGAGAAGGGTCTTTATGGGTTGTATTCAATTCAGAAACA 70
DB 1 ATGAGAGATAGGTAAAGATAGGGGAAGGATCTTATGGCATTTGTTTAAATGTCGGAACC 60
QY 71 AAACCTCTGGACAAGTAGTAGCTTTTAAATAAATTTTGGAAATCTGAAGATGATCCTGTTG 130
DB 61 GTGACTCGGGACAAATTTGTGGCATCAAAAATTTTGGAGTCAGAGAGATGACCCGCTGA 120
QY 131 TTAAGAAATAGCACAATAGAGAAATACGTATGTTGAAGCAATTAACATCAATCTTG 190
DB 121 TCAAGAAATCGCGCTGAGGGAATCAGAAATGTTAAAGCAATTAACACAAATAACTTGG 180
QY 191 TGAACCTCATCGAGGTGTTTCAAGAGAAAAGGAAAATGCAATTTAGTTTGAATACCTGTG 250
DB 181 TGAACCTGATGAAGTTTTCGTCGAAAACGCAATACATCTTGTGTCGAGTATGTG 240
QY 251 ATCATACACTTTTAAATGAGCTGGAAAAGAAACCCCAATGGAGTTGCTGATGAGTGATCA 310
DB 241 ATCATACAGTTTAAATGAATTTGACAAACACATGAGAGAGTACCAGAACATCATGTTA 300
QY 311 AAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCTATATACATACTGTTATTC 370
DB 301 AACGTAATGTTTGGCAAGTTCTTCAAGCTGTTCACTTTGCCATCAGCATATGTATAC 360
QY 371 ACAGAGATATAAAACCTGAAAATATTCTAATAACTAAGCAAGGAATAATCAAGATTGTG 430
DB 361 ATCGTGATGTAAACCCAGAGAACATATTAACTACTAAGCAAGGTGTTATCAAGTTGTG 420
QY 431 ACTTCGGGTTTGACAAATTTCTGA---TTCCAGGAGATGCTTACACCGATTATGTAGCTA 487
DB 421 ACTTTGGATTGCAAGAAATTTTAACTGGTCTCTGGCCAGCAATTACACCGACTATGTGCCA 480
QY 488 CGAGATGGTACCGAGCTCTGAACTTCTTGTGGGAGATGCTCAGTATGTTCTTCACTCG 547
DB 481 CTAGGTGTTACAGAGCACCTGAGCTCTCTCGTGGGGACACTCAGTATGCGCCCTCCAGTGG 540

QY 548 ATATATGGGCTATTGGTTGTGTTTTTGCAGAGCTCCTGACAGGCCAGCCACTGTGGCC 605
DB 541 ATGTGTGGGCAATAGTTTGGTATTTGCCGAGTTGTTGTGCGGACAGGCATTATGGCC 598
RESULT 11
BP016186 597 bp mRNA linear EST 15-MAR-2002
LOCUS BP016186
DEFINITION BP016186 Nori Satoh unpublished cDNA library, young adult Ciona
intestinalis cDNA clone ciad64e08 5', mRNA sequence.
ACCESSION BP016186
VERSION BP016186.1 GI:19507663
KEYWORDS EST.
SOURCE Ciona intestinalis
ORGANISM Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Cionidae; Ciona.
REFERENCE 1 (bases 1 to 597)
AUTHORS Satoh,N., Satou,Y., Kohara,Y. and Shin-i,T.
TITLE Expressed genes in Ciona intestinalis
JOURNAL Unpublished
COMMENT Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoheascidian.zool.kyoto-u.ac.jp.
FEATURES
source
1..597
/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="ciad64e08"
/tissue_type="whole animal"
/dev_stage="young adult"
/clone_lib="Nori Satoh unpublished cDNA library, young
adult"
BASE COUNT 184 a 100 c 139 g 173 t 1 others
ORIGIN
Query Match: 30.6%; Score 289.2; DB 12; Length 597;
Best Local Similarity 71.0%; Pred. No. 1.4e-53;
Matches 397; Conservative 0; Mismatches 159; Indels 3; Gaps 1;
QY 2 TGGAAAAGTATGAAAATATTAGCTAAGACTGGAGAAGGGTCTTATGGGTTGTATTCAAT 61
DB 38 TGGAGAAATATGAGAAGATAGGTAAGATAGGGGAAGGATCTATGGCATTTGTGTTAAAT 97
QY 62 GCAGAAACAAACCTCTGGACAAGTAGTAGCTGTTAAAAAATTTGTGGAATCTGAAGATG 121
DB 98 GTCGGAACCGTGACTCGGACAAATTTGTGGCCATCAAAAATTTGTGAGTCAGAAGATG 157
QY 122 ATCCTGTTGTTAAGAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACACATC 181
DB 158 ACCCGCTGATCAAGAAAATCGCGCTGAGGGAATCAGAAATGTTAAAGCAATTTGAACACA 217
QY 182 CAAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAGSAAAATGCATTTAGTTTGTG 241
DB 218 ATAACCTTGGTGAACCTGATTGAAGTTTTCGTCGAAAACGCAATTTACATCTTGTGTCG 277
QY 242 AATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCCAATGGAGTTGCTGATG 301
DB 278 AGTATTGTGATCATACAGTTTAAATGAATTTGGACAAACACATGAGAGGAGTACCAGAAC 337
QY 302 GAGTGATCAAAAGCGTATTATGCGAAACACTTCAAGCTCTTAAATTTCTGTATATACATA 361
DB 338 ATCATGTTAAACCGTATTGTTTGGCAAGTTCTTCAAGCTGTTCACTTTTGCCATCAGCATA 397
QY 362 ACTGTATTACAGAGATATAAAACCTGAAAATATTCTAATAACTAAGCAAGGAATATCA 421
DB 398 ATTGTATACATCCTGTGATGTAACACCCAGAGAACATATTAACTACTAAGCAAGGTGTTATCA 457


```
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:2236229"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Sugano mouse kidney mkia"
/note="Organ: kidney; Vector: pME18S-FL3; Site 1: DraIII
(CACTGTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGGCCCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor [TGTGGCCTACTGG], digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTCTGCTCTAAAGCTGCG and 3' end
primer CGACCTGGAGCTCGAGCACA."
BASE COUNT      202 a   179 c   191 g   157 t      2 others
ORIGIN

Query Match      27.4%; Score 258.6; DB 9; Length 731;
Best Local Similarity 66.9%; Pred. No. 8.7e-47;
Matches 382; Conservative 0; Mismatches 186; Indels 3; Gaps 1;

QY 5 AAAAGTATGAAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAAAATGCA 64
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 156 AAATATGAAAAAAATTTGGAAGATTGGAGAGGGCTCTATGGGTAGTGTCAAGTSCA 215

QY 65 GAAACAAAACCTCTGGACAAGTAGTAGCTGTTTAAAAAATTTTGGAAATCTGAGATGATC 124
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 216 GAAACAGGGACACGGGTCAGATCGTGGCCATCAAGAGGTTTCTGGAAACCGAGATGACC 275

QY 125 CTGTTGTTAAGAAAAATAGCATAAGAGAAATACGTATGTTGAAGCAATTAAACATCCAA 184
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 276 CTGTCTAAGAAAAATCGCCCTCGAGAAATCCGCATGCTCAAGCAACTCAAGCACCCCA 335

QY 185 ATCTTGTAACCTCATCGAGGTGTTTCAGGAGAAAAAGAAAAATGCATTTAGTTTGAAT 244
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 336 ACCTGGTCAACCTCTCTGGAAGTCTTCGGAGGAAAGCGGAGGCTTCACCTGGTGTTCGAGT 395

QY 245 ACTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGGAGTTGCTGATGGAG 304
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 396 ACTGGACACACACGGTGCTTCCAGAGCTGGATCGGTATCAGAGGGGGGTACCAAGAGCCTC 455

QY 305 TGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCTATACATAACT 364
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 456 TCGTGAAGAACATAAATTTGGCAGACACTGCAGGCTGTAATTTCTGCCATFAAACATAACT 515

QY 365 GTATTACAGAGATATAAAACCTGAAAAATATTCTAATAACTAAGCAAGGATAATCAAGA 424
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 516 GCATACAGAGACGTCGTAAGCCGGAAAAATATTCTCATCAACCAAAACAGTCAGGCATTAGC 575

QY 425 TTTGTGACTTCGGGTTTGACAAATTTCTGATT---CCAGGAGATGCCTACACCGATTATG 481
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 576 TCTGTGACTTTGGGGTCGCACCGCTCCTCACTGGACCANGTGACTACTACACAGACTACN 635

QY 482 TAGCTACAGATGGTACCGAGCTCTGAACTTCTTGTGGAGATACCTCAGTATGGTTCTT 541
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 636 GTGCCCCCGGTGGTACTGCTCACCCGAGCTGCTAGTTGGAGACACCGCAGTATGGTCCC 695

QY 542 CAGTCGATATATGGGCTATTGGTTGTGTTTT 572
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 696 CTGTAGATGTCTGGGCAATTGGGCTGTGTTT 726
```

```
RESULT 14
CA334835/c
LOCUS
DEFINITION
CA334835      645 bp      mRNA      linear      EST C4-NOV-2002
NISC_lc01h12.y1 COGENE 8.5 EPT Homo sapiens cDNA clone
IMAGE:560587C 5', mRNA sequence.
```

```
ACCESSION      CA334835
VERSION        CA334835.1      GI:24552933
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE        1 (bases 1 to 645)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
Tumor Gene Index
Unpublished
CONTACT      Robert Strausberg, Ph.D.
EMAIL      cgapbs-remail.nih.gov
CDNA LIBRARY Preparation:
cDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC)
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
info@image.llnl.gov
Plate: LLAM12404 row: 0 column: 23
Seq primer: M13RPI reverse primer (ABI).
Location/Qualifiers
1..645
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5605870"
/tissue_type="posterior tongue (pooled)"
/dev_stage="embryo, 8 weeks postconception"
/lab_host="DH10B"
/clone_lib="COGENE 8.5 EPT"
/note="Vector: pAMP1; cDNA primed using oligo-dT primer,
directionally cloned into UDG sites of pAMP1. Size
selected for insert sizes ranging from 0.2-2.5 kb.
Normalized to Cot10. Primary library, non-amplified.
Library constructed by M. Lovett. For more information on
this library, please contact R. Tidwell (Washington
University) or visit the COGENE website at
http://hg.wustl.edu/COGENE/."
BASE COUNT      153 a   158 c   147 g   187 t
ORIGIN

Query Match      27.0%; Score 255; DB 14; Length 645;
Best Local Similarity 63.5%; Pred. No. 5.3e-46;
Matches 406; Conservative 0; Mismatches 230; Indels 3; Gaps 1;

QY 275 AAGAAACCCAAATGAGTTGCTGATGGAGTGATCAAAAGCGTATTATGGCAACACTTC 334
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 645 ACAGATACCAAGAGGGGTACCAGAACATCTCGTGAAGAGCATAACTTGGCAGACACTGC 586

QY 335 AAGCTCTTAATTTCTGTATATACATAACTGTATTTCACAGAGATATAAAACCTGAAATA 394
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 585 AAGCTGTAATTTTGGCCATAAACACAAATTCATACATAGAGACGTGAAGCCAGAAATA 526

QY 395 TTCTAATAACTAAGCAAGGAATAATCAAGATTGTGACTTCGGGTTTGCACAAATTCCTGA 454
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 525 TCCTCATCACGAAACATTCCTGATTAAAGCTTTGTGACTTTTGGATTGCTCGGCTTTTGA 466

QY 455 TT---CCAGGAGATGCTACACCGATTATGTAGCTACGAGATGGTACCGAGCTCCTGAAC 511
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 465 CTGGACCGAGTGACTACTATACAGACTACCTGGCTACCCAGGTGGTACCGCTCCCTGAGC 406

QY 512 TTCTTGTGGAGATACTCAGTATGGTTCTTCAGTCGATATATGGGCTATTTGGTTGTGTTT 571
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 405 TGCTGGTGGGGACACGAGTACCGCCCCCGGTGGATGTTTGGCAATTGGCTGTGTCT 346

QY 572 TTGCAGAGCTCCTGACAGGCCAGCCACTGTGGCCTGGAAAAATCAGATGTGGACCACTTT 631
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 345 TTGCTGAGCTGCTGTGAGGAGTGCCTCTGTGGCCAGGAAAAATCGGATGTGGATCAGCTGT 286

QY 632 ATCTGATAATCAGAACACTAGGAAAAATTAATCCCAAGACATCAATCAATCTTTAAAGTA 691
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2003, 16:54:37 ; Search time 73.9079 Seconds
(without alignments)
6216.928 Million cell updates/sec

Title: US-09-671-050-9
Perfect score: 1041
Sequence: 1 atggaaaagtatgaaaatt.....aggtacttcggtcaaaagt 1041

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220591566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCITUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	251.8	24.2	1701	4 US-09-411-628-5	Sequence 5, Appli
2	251.8	24.2	3080	4 US-09-411-628-3	Sequence 3, Appli
3	163.8	15.7	903	2 US-08-874-347-9	Sequence 9, Appli
4	163.8	15.7	903	3 US-09-093-522-9	Sequence 1, Appli
5	162	15.6	1476	2 US-08-969-106-1	Sequence 1, Appli
6	162	15.6	1476	4 US-09-338-125-1	Sequence 1, Appli
7	162	15.6	1476	4 US-09-266-225D-13	Sequence 13, Appl
8	162	15.6	1635	4 US-09-417-197-112	Sequence 112, App
9	162	15.6	1635	4 US-09-417-197-114	Sequence 114, App
10	154.2	14.8	1050	4 US-09-220-132-3	Sequence 3, Appli
11	134.8	12.9	1825	4 US-09-620-312D-313	Sequence 313, App
12	133.2	12.8	1089	1 US-08-154-915-1	Sequence 1, Appli
13	133.2	12.8	1089	2 US-08-464-517-37	Sequence 37, Appli
14	133.2	12.8	1089	2 US-08-246-361A-37	Sequence 37, Appli
15	133.2	12.8	1089	3 US-08-463-772-37	Sequence 37, Appli
16	133.2	12.8	1089	5 PCT-US93-09945-1	Sequence 1, Appli
17	128	12.3	1002	1 US-08-463-090B-3	Sequence 3, Appli
18	116	11.1	1070	1 US-08-463-090B-5	Sequence 5, Appli
19	112	10.8	1308	4 US-09-801-861-4	Sequence 4, Appli
20	112	10.8	2203	4 US-09-801-861-1	Sequence 1, Appli
21	103	9.9	2747	2 US-08-874-347-1	Sequence 1, Appli
22	103	9.9	2747	3 US-09-093-522-1	Sequence 1, Appli
23	100.6	9.7	2467	4 US-09-206-344A-3	Sequence 3, Appli
24	96.2	9.2	1673	3 US-09-347-801-15	Sequence 15, Appli
25	94.2	9.0	1272	4 US-09-206-344A-1	Sequence 1, Appli
26	94.2	9.0	1549	4 US-09-206-344A-13	Sequence 13, Appli
27	89.6	8.6	90541	4 US-09-759-359A-3	Sequence 3, Appli

c	28	87.4	8.4	6340	1	US-08-187-785-3	Sequence 3, Appli
	29	87.4	8.4	87350	3	US-08-781-891-79	Sequence 79, Appl
	30	87.4	8.4	87350	4	US-09-618-166-79	Sequence 79, Appl
	31	87.4	8.4	87543	4	US-09-791-211-3	Sequence 3, Appli
	32	87.4	8.4	99500	4	US-09-798-036-10	Sequence 10, Appl
	33	87.2	8.4	1946	4	US-09-620-312D-462	Sequence 462, App
	34	87.2	8.4	70000	4	US-09-851-896-3	Sequence 3, Appli
c	35	87	8.4	8779	2	US-08-750-703-4	Sequence 4, Appli
	36	86.6	8.3	2624	1	US-08-032-382B-1	Sequence 1, Appli
	37	86.4	8.3	66804	4	US-09-740-041-3	Sequence 3, Appli
	38	86	8.3	72604	4	US-09-268-992-7	Sequence 7, Appli
	39	86	8.3	72604	4	US-09-657-474-7	Sequence 7, Appli
	40	85.6	8.2	604	4	US-09-659-845A-22	Sequence 22, Appl
c	41	85.6	8.2	16595	4	US-09-146-053-7	Sequence 7, Appli
c	42	85.4	8.2	7216	2	US-08-750-703-3	Sequence 3, Appli
	43	85.4	8.2	112132	4	US-09-741-150-3	Sequence 3, Appli
c	44	85.2	8.2	18853	4	US-09-820-005-3	Sequence 3, Appli
	45	84.6	8.1	1249	4	US-09-220-132-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-09-411-628-5
; Sequence 5, Application US/09411628
; Patent No. 6428994
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; TITLE OF INVENTION: CDNA, GENOMIC, AND PREDICTED PROTEIN
; TITLE OF INVENTION: SEQUENCES OF LEARNING-INDUCED KINASES
; FILE REFERENCE: 13761-707
; CURRENT APPLICATION NUMBER: US/09/411,628
; CURRENT FILING DATE: 1999-10-01
; EARLIER APPLICATION NUMBER: US 60/102,906
; EARLIER FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1701
; TYPE: DNA
; ORGANISM: Oryctolagus cuniculus
US-09-411-628-5

Query Match	24.2%	Score 251.8;	DB 4;	Length 1701;
Best Local Similarity	62.9%	Pred. No. 9.1e-61;		
Matches 407; Conservative	0;	Mismatches 237;	Indels 3;	Gaps 1;
Qy	1	ATGGAAAAGTATGAAAATTTAGCTAAGACTGGAGAAGGGTCTTATGGGGTTGTATTCAA	60	
Dc	1	ATGGAAAATATGAGAACCTTGGATTGGTTGGAGAAGGGAGTTATGGAATGGTGATGAAG	60	
Qy	61	TGCAGAAACAAACCTCTCGACAAGTAGTAGCTGTTAAAAAATTTGTGGAATCTGAAGAT	120	
Dc	61	TGTAGGAATAAGACATAGTGGAGAATTTGGCCCATCAAGAAGTTCTTAGAAAAGTGATGAT	120	
Qy	121	GATCCTGTTGTTAAGAAAATAGCACTAAGAGAAAATACGTATGTTGAAGCAATTAAAACAT	180	
Dc	121	GACAAAATGGTTAAAAAATTTGCTATGCGAGAAAATCAAGTTACTAAAGCAACTGAGGCAT	180	
Qy	181	CCAAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAGGAAAATGCAATTTAGTTTTT	240	
Dc	181	GAAAATTTGGTGAATCTGTTGGAGGTGTGTAAAAAANAACCGATGGTACCTAGTCTTT	240	
Qy	241	GAATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT	300	
Dc	241	GAATTTGTTGACCACACGATTTCTTGATGACTTGGAACTCTTTCCAAATGGACTAGATGAC	300	
Qy	301	GGAGTGAATCAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGTATATACAT	360	
Dc	301	CAAGTAGTTCAAAAGTATTGTTTTCAGATTATTAAATGGAATGGATTTGTACAGTCAC	360	
Qy	361	AACTGTATTACAGAGATATAAAACCTGAAAATATTCTAATACTAAGCAAGGAATAATC	420	

Db 361 AATATCATACATAGAGATATAAGCCAGAGAAATATATTGGTCTCCAGTCTGGCGTTGTC 420
Qy 421 AAGATTTGTGACTTCGGGTTTGCACAAATTCG---ATTCCAGGAGATGCCCTACACCGAT 477
Db 421 AAGTTATGTGATTTTGGATTTCACGGACACTGGCAGCTCCCGGAGAGGTTTACACTGAT 480
Qy 478 TATGTAGCTACGAGATGTACCGAGCTCCTGAACTTCTTGTGGGAGATACCTCAGTATGGT 537
Db 481 TATGTGGCAACTCGATGGTACAGAGCTCCAGAACTACTGGTTGGTGTGATGTCAGTATGGC 540
Qy 538 TCTTCAGTCGATATATGGGCTATTGGTTGTGTTTTTGCAGAGCTCCTGACAGGCCAGCCA 597
Db 541 AAAGCTGTGGATGTGGGCCCAATTGGTTGTCTGGTAACTGAAATGCTCATGGGGGAACCC 600
Qy 598 CTGTGGCCTGGAAATCAGATGTGGACCAACTTTATCTGATAATCAG 644
Db 601 CTGTTTCTCGAGACTCTGATATTGATCAGCTTTATCTTATTATGAG 647

RESULT 2

US-09-411-628-3
; Sequence 3, Application US/09411628
; Patent No. 6428994
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; TITLE OF INVENTION: CDNA, GENOMIC, AND PREDICTED PROTEIN
; TITLE OF INVENTION: SEQUENCES OF LEARNING-INDUCED KINASES
; FILE REFERENCE: 13761-707
; CURRENT APPLICATION NUMBER: US/09/411,628
; CURRENT FILING DATE: 1999-10-01
; EARLIER APPLICATION NUMBER: US 60/102,906
; EARLIER FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3080
; TYPE: DNA
; ORGANISM: Orcytolagus cuniculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (249)...(1949)

US-09-411-628-3

Query Match 24.2%; Score 251.8; DB 4; Length 3080;
Best Local Similarity 62.9%; Pred. No. 1.2e-60;
Matches 407; Conservative 0; Mismatches 237; Indels 3; Gaps 1;

Qy 1 ATGGAAGAAGTATGAAAATTAGCTAAGACTGGAGAGGGTCTTTATGGGTTGTATTCAA 60
Db 249 ATGGAAGAATATGAGAACCTTGGATTGGTTGGAGAGGGAGTTATGGAATGGTATGAAG 308
Qy 61 TGCAGAAACAAACCTCTGCACAAGTAGTAGCTGTAAATAATTTGTGGAATCTGAAGAT 120
Db 309 TGTAGGAATAAAGATAGTGAAGAATTGTGGCCATCAAGAAGTTCTTAGAAAGTGATGAT 368
Qy 121 GATCCTGTTGTTAAGAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTTAAACAT 180
Db 369 GACAAAATGGTTAAAAAATTGCTATGCGAGAAATCAAGTTACTAAAGCAACTGAGGCAT 428
Qy 181 CCAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAAGGAAATGCAATTAGTTTT 240
Db 429 GAAAATTGGTGAATCTGTTGAGGTGTGTAAAAAATAAATAAAGATGCTAGTCTTT 488
Qy 241 GAATACCTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT 300
Db 489 GAATTTCTTGACCACACCGATTCTTGATGACTTGGAACTCTTTCCAAATGGACTAGATGAC 548
Qy 301 GGAGTGATCAAAAGCGTATTATTGGCAACACTTCAAGCTCTTAAATTTCTGTCATATACAT 360
Db 549 CAAGTAGTCAAAAGTATTTGTTTCAGATTATTAAATGGAATTTGGAATTTTGTACAGTCA 608
Qy 361 AACTGTATTACAGAGATATATAAACCTGAAATATTTCTTAATAACTAAGCAAGGAATAATC 420

Db 609 AATATCATACATAGAGATATAAGCCAGAGAAATATATTGGTCTCCAGTCTGGCGTTGTC 668
Qy 421 AAGATTTGTGACTTCGGGTTTGCACAAATTCG---ATTCCAGGAGATGCCCTACACCGAT 477
Db 669 AAGTTATGTGATTTTGGATTTCACGGACACTGGCAGCTCCCGGAGAGGTTTACACTGAT 728
Qy 478 TATGTAGCTACGAGATGTACCCAGAGCTCCTGAACTTCTTGTGGGAGATACCTCAGTATGGT 537
Db 729 TATGTGGCAACTCGATGGTACAGAGCTCCAGAACTACTGGTTGGTGTGATGTCAGTATGGC 788
Qy 538 TCTTCAGTCGATATATGGGCTATTGGTTGTGTTTTTGCAGAGCTCCTGACAGGCCAGCCA 597
Db 789 AAAGCTGTGGATGTGTGGGCCAATTGGTTGTCTGGTAACTGAAATGCTCATGGGGGAACCC 848
Qy 598 CTGTGGCCTGGAAATCAGATGTGGACCAACTTTATCTGATAATCAG 644
Db 849 CTGTTTCTCGAGACTCTGATATTGATCAGCTTTATCTTATTATGAG 895

RESULT 3

US-08-874-347-9
; Sequence 9, Application US/08874347
; Patent No. 5863741
; GENERAL INFORMATION:
; APPLICANT: Limper, Andrew H.
; APPLICANT: Leof, Edward B.
; APPLICANT: Thomas, Charles F.
; APPLICANT: Gustafson, Michael P.
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
; TITLE OF INVENTION: CARINI
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C., P.A.
; STREET: 60 South Sixth Street, Suite 3300
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/874,347
; FILING DATE: 13-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ellinger, Mark S.
; REGISTRATION NUMBER: 34,812
; REFERENCE/DOCKET NUMBER: 07039/055001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-335-5070
; TELEFAX: 612-288-9696
; TELEX:
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 903 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 1...900
; OTHER INFORMATION:

US-08-874-347-9

Query Match 15.7%; Score 163.8; DB 2; Length 903;
Best Local Similarity 55.2%; Pred. No. 2.7e-36;
Matches 365; Conservative 0; Mismatches 287; Indels 9; Gaps 2;

```
QY 1 ATGGAAAAGTATGAAAAAATTAGCTAAGACTGGAGAAAGGCTCTATGGGTTGTAATCAA 60
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1 ATGGAGCAATATCAGAGGTTAGAGAAGATTGGAGAAGGAACCTATGGAGTTGTTTATAA 60
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 61 TGCAGAAACAAACCTCTGGACACAGTAGTAGCTGTTAAAAAATTTGTGGAATCTGAAGAT 120
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 61 GCAAAGGATCTTGAAGTGGTACAATTTAGCTCTTAAGAAAAATCCGATTAGAAGCAGAA 120
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 121 GATCCTGTTGTTAAGAAAAATAGCCTACAGAGAAATACGTATGTTGAAGCAATTAACAT 180
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 121 GATGAGGAGTTCTTAGTACAGCAATTCGTGAGATATCACCTTTTSAAGAGATGCACAAT 180
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 181 CCAAAATCTTGTGAACCTCATCGAGGTGTTCAAGAGAAAAAGGAAATGCATTTAGTTT 240
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 181 GATAATGTTGAAGACTTTTGAATATAAATCATCAAGAGTCACGTTTATATCTTGTTTT 240
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 241 GAATACCTGTATACATACCTTTTAAATAGAGTGGAAAGAAACCCAAATG-----GAGTT 294
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 241 GAATTTCTTGATCTGATTTAAAAAAGTATATGAATAGTATTTCCAAAAGGACATGATGCTT 300
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 295 GCTGATGAGTGATCAAAAGCGTATTATGCGAACAACCTTCAAGSTCTTAATTTCTGTCAT 354
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 301 GGTGCAGAAATGATTAAGAAAGTTTATGTCAACAATTTGTATCAGTGTATAATATTGTCAT 360
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 355 ATACATAACTGTATTCACAGAGATATAAAACCTGAAAAATATTCTAATAACTAAGCAAGGA 414
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 361 TCTCATCGTATTCCTCATCGTACTTGAACCCACAAATCTTCTATAGATCGAGAGGA 420
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 415 ATAATCAAGATTGTGACTTCGGGTTTGCACAAA---TTCTGATTCAGGAGATGCCTAC 471
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 421 AATCTTAAATTAGCAGATTTTGGGCTTGCAGGGCGTTTGGTGTCCATTGCGTGGTTAT 480
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 472 ACCGATTATGTAGCTACGAGATGGTACCGAGTCTCTGAACCTTCTTGTGGGAGATACCTCAG 531
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 481 ACTCATGAAGTTGTTACACTTTGGTATCGTGTCCAGAAAGTTCTTTTAGGTGGTCGACAA 540
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 532 TATGTTCTTCAGTCGATATATGGGCTATTGGTGTGTTTTCAGAGACTCCTGACAGGC 591
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 541 TATGCAACAGCGCTGATATATGGAGCATTTGATGTATTTTTCAGAAATGGCTACAAA 600
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 592 CAGCCACTGTGGCTGGAAAAATCAGATGTGGACCAACTTTTATCTGATAATCAGAACACTA 651
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 601 AAGCCATTATTTCCAGGTGATCTCTGAATGTGATGAATATTTAGANATATTTAGATATTA 660
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 652 G 652
  |
Db 661 G 661
```

RESULT 4

```
US-09-093-522-9
; Sequence 9, Application US/09093522
; Patent No. 6015700
; GENERAL INFORMATION:
; APPLICANT: Limper, Andrew H.
; APPLICANT: Leof, Edward B.
; APPLICANT: Thomas, Charles F.
; APPLICANT: Gustafson, Michael P.
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
; TITLE OF INVENTION: CARINI
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C., P.A.
; STREET: 60 South Sixth Street, Suite 3300
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/093,522
; FILING DATE: 08-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/874,347
; FILING DATE: 13-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ellinger, Mark S.
; REGISTRATION NUMBER: 34,812
; REFERENCE/DOCKET NUMBER: 07039/055002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-335-5070
; TELEFAX: 612-288-9696
; TELEX:
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 903 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 1..900
; OTHER INFORMATION:
;
US-09-093-522-9
```

```
Query Match 15.7%; Score 163.8; DB 3; Length 903;
Best Local Similarity 55.2%; Pred. No. 2,7e-36;
Matches 365; Conservative 0; Mismatches 287; Indels 9; Gaps 2;
```

```
QY 1 ATGGAAAAGTATGAAAAAATTAGCTAAGACTGGAGAAAGGCTCTTATGGGTTGTAATCAA 60
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1 ATGGAGCAATATCAGAGGTTAGAGAAGATTGGAGAAGGAACCTATGGAGTTGTTTATAA 60
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 61 TGCAGAAACAAACCTCTGGACACAGTAGTAGCTTAAAAAATTTGTGGAATCTGAAGAT 120
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 61 GCAAAGGATCTTGAAGTGGTACAATTTAGCTCTTAAGAAAAATCCGATTAGAAGCAGAA 120
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 121 GATCCTGTTGTTAAGAAAAATAGCCTACAGAGAAATACGTATGTTGAAGCAATTAACAT 180
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 121 GATGAGGAGTTCTTAGTACAGCAATTCGTGAGATATCACCTTTTGAAGAGATGCACAAT 180
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 181 CCAAAATCTTGTGAACCTCATCGAGGTGTTCAAGAGAAAAAGGAAATGCATTTAGTTT 240
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 181 GATAATGTTGAAGACTTTTGAATATAAATCATCAAGAGTCACGTTTATATCTTGTTTT 240
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 241 GAATACCTGTATACATACCTTTTAAATAGAGTGGAAAGAAACCCAAATG-----GAGTT 294
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 241 GAATTTCTTGATCTGATTTAAAAAAGTATATGAATAGTATTTCCAAAAGGACATGATGCTT 300
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 295 GCTGATGAGTGATCAAAAGCGTATTATGCGAACAACCTTCAAGCTCTTAATTTCTGTCAT 354
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 301 GGTGCAGAAATGATTAAGAAAGTTTATGTCAACAATTTGTATCAGGTTTAAATATTGTCAT 360
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 355 ATACATAACTGTATTCACAGAGATATAAAACCTGAAAAATATTCTAATAACTAAGCAAGGA 414
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 361 TCTCATCGTATTCCTCATCGTACTTGAACCCACAAATCTTCTATAGATCGAGAGGA 420
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 415 ATAATCAAGATTGTGACTTCGGGTTTGCACAAA---TTCTGATTCAGGAGATGCCTAC 471
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 421 AATCTTAAATTAGCAGATTTTGGGCTTGCAGGGCGTTTGGTGTCCATTGCGTGGTTAT 480
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 472 ACCGATTATGTAGCTACGAGATGGTACCGAGTCTCTGAACCTTCTTGTGGGAGATACCTCAG 531
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 481 ACTCATGAAGTTGTTACACTTTGGTATCGTGTCCAGAAAGTTCTTTTAGGTGGTCGACAA 540
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 532 TATGTTCTTCAGTCGATATATGGGCTATTGGTGTGTTTTCAGAGACTCCTGACAGGC 591
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 541 TATGCAACAGCGCTTGTATATGGAGCATTTGATGTATTTTTCAGAAATGGCTACAAA 600
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 592 CAGCCACTGTGGCTGGAAAAATCAGATGTGGACCAACTTTTATCTGATAATCAGAACACTA 651
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
```


Db 601 AAGCCATTATTCCAGGTGATTCTGAAATTGATGAAATATTTAGAAATATTAGAAATATTA 660
QY 652 G 652
Db 661 G 661

RESULT 5
US-08-969-106-1
; Sequence 1, Application US/08969106
; Patent No. 5986055
; GENERAL INFORMATION:
; APPLICANT: Yang, M.
; APPLICANT: Nandabalan, K.
; APPLICANT: Schulz, V.
; TITLE OF INVENTION: CDK2 INTERACTIONS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.3
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/969,106
; FILING DATE: 13-NOV-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7934-057
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1476 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 1..894
; OTHER INFORMATION:
US-08-969-106-1

Query Match 15.6%; Score 162; DB 2; Length 1476;
Best Local Similarity 55.0%; Pred. No. 1.le-35;
Matches 362; Conservative 0; Mismatches 290; Indels 6; Gaps 2;
QY 1 ATGGAAGACTATGAAATAATTAGCTAAGACTGGAGAGGGCTTTATGGGTTGTATTCAA 60
Db 1 ATGGAGAACTTCCAAAGGTTGGAAGAAGATCGGAGAGGSCACGTACGGAGTTGTGTACAA 60
QY 61 TGCAGAAACAAACCTCTGGACACTAGTAGCTGTTTAAATAATTTGTGGAATCTGAAGAT 120
Db 61 GCCAGAAACAAAGTTGACGGGAGAGGTTGGTGGCTTAAGAAATCCGGCTGGACACTGAG 120
QY 121 GATCCTGTGTTAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAAACAT 190
Db 121 ACTGAGGTTGTGCCCCAGTACTGCCATCCGAGAGATCTCTCTGCTTAAGGAGCTTAACCAT 180
QY 181 CCAATCTTGTGAACCTCATCGAGGTGTTCCAGGAGAAAGGAAATGCAATTAGTTTTT 240

Db 181 CCTAATATGTCAAGCTGCTGGATGTCTATTCACACAGAAATAAACTCTACCTGGTTTTT 240
QY 241 GAATACCTGTGATCATACACTTTTAAATGAGCTGGA---AAGAAACCCAAATGGAGTTGCT 297
Db 241 GAATTTCTGCACCAAGATCTCAAGAAATTCATGGATGCTCTCTCACTGGCATCTCT 300
QY 238 GATGGAGTGATCAAAAGCGTATTATGGCAACACACTTCAAGCTCTTAATTTCTGTCTATATA 357
Db 301 CTTCCTCTCATCAAGAGCTATCTGTTCCAGCTGCTCCAGGGCTAGCTTTCTGCTCATTTCT 360
QY 358 CATAACTGTATTACAGAGATATATAAACCTGAAATAATTTCTAATACTAAGCAAGGAATA 417
Db 361 CATCGGCTCTCCACCGAGACCTTAAACCTCAGAAATCTGCTTATTAAACACAGAGGGGCC 420
QY 418 ATCAAGATTGTGACTTCGGGTTTGCACAAATTCCTG---ATTCCAGGAGATGCCCTACACC 474
Db 421 ATCAAGCTACAGAGACTTTGGACTAGCCAGAGCTTTTGGAGTCCCTGTTCTGACTTACACC 480
QY 475 GATTATGTAGCTACGAGATGTTACCGAGCTCCTGAACCTTTGTTGGGAGATACCTCAGTAT 534
Db 481 CATGAGGTGTTGACCTGTGTGTTACCGAGCTCCTGAAATCTCTCTGGGCTCGAAATATTAT 540
QY 535 GGTTCTTCACTCGATATATGGGCTATTGTTGTTGTTTTCAGAGAGCTCCTGACAGGCCAG 594
Db 541 TCCACAGCTGTGGACACTGTGGAGCTGGGCTGCATCTTTGCTGAGATGGTGACTCGCCGG 600
QY 595 CCACTGTGGCTGGAATAATCAGATGTGGACCAACTTTA*CTGATAATCAGAAACACTAG 652
Db 601 GCCCTGTTCCCTGGAGATTCTGAGATTGACCAGCTCTTCCGGATCTTTTCGGACTCTGG 658

RESULT 6
US-09-338-125-1
; Sequence 1, Application US/09338125
; Patent No. 6521412
; GENERAL INFORMATION:
; APPLICANT: Yang, M.
; APPLICANT: Nandabalan, K.
; APPLICANT: Schulz, V.
; TITLE OF INVENTION: CDK2 INTERACTIONS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/338,125
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/969,106
; FILING DATE: 13-NOV-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7934-057
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1476 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

```

/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA
/ FEATURE:
/   NAME/KEY: Coding Sequence
/   LOCATION: 1...894
/   OTHER INFORMATION:
/ US-09-338-125-1

```

Query Match	15.6%	Score 162;	DB 4;	Length 1476;
Best Local Similarity	55.0%	Pred. No. 1.le-35;		
Matches 362;	Conservative 0;	Mismatches 290;	Indels 6;	Gaps 2;
QY	1	ATGGAAAGTATGAAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAA	60	
DB	1	ATGGAGAACTTCCAAAGAGTGGAAAGATCGGAGAGGGCAGCTACGGAGTTGTGTACAA	60	
QY	61	TGCAGAAACAAAACCTCTGGACAAGTAGTAGCTGTTAAATAATTTGTGGAATCTGAAGAT	120	
DB	61	GCCAGAAACAAGTTGACGGGAGAGGTGGTGGCGCTTAAGAAAAATCCGCCCTGSACACTGAG	120	
QY	121	GATCCTGTTGTTAAGAAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAATAACAT	180	
DB	121	ACTGAGGGTGTGCCAGTACTGCCATCCGAGAGATCTCTCTGCTTAAGGAGCTTAACCAT	180	
QY	181	CCAAATCTTGTAACCTCATCGAGGTGTTTCAGGAGAAAAAGCAAAATGCATTTAGTTTTT	240	
DB	181	CCTAATATTGTCAAGCTGCTGGAATGTCATTACACACAGAAAAATAAACTCTACCTGGTTTTT	240	
QY	241	GAATACTGTGATCATACACTTTTAAATGAGCTGSA---AAGAAACCCAAATGGAGTTGCT	297	
DB	241	GAATTTCTGCACCAAGATCTCAAGAAATTCATGGATGCCCTCTGCTCTCACTGGCATTCT	300	
QY	298	GATGGAGTGATCAAAAGCGTATTATGGCAACACITTCAGCTCTTAATTTCTGTCAATA	357	
DB	301	CTTCCCTCATCAAGAGCTATCTGTTCCAGCTGCTCCAGGGCCTAGCTTTCTGCCATTCT	360	
QY	358	CATAACTGTATTACAGAGATATAAAACCTGAAAAATATCTTAATAACTAAGCAAGGAATA	417	
DB	361	CATCGGGTCTCCACCGAGACCTTAACCTCAGAAATCTGCTTATTAACACAGAGGGGCC	420	
QY	418	ATCAAGATTGTGACTTCGGGTTTGCAAAATTCG---ATTCAGGAGATGCCTACACC	474	
DB	421	ATCAAGCTAGCAGACTTTGGACTAGCCAGAGCTTTTGGACTCCCTGTTCTGTAATTACAC	480	
QY	475	GATTATGTAGCTACGAGATGGTACCGAGCTCCTGAACTTTCTTGTGGGAGATACCTCAGTAT	534	
DB	481	CAIGAGGTGTGACCCCTGTGGTACCGAGCTCCTGAAATCCTCTGGGCTCGAAATATTAT	540	
QY	535	GGTCTCTCAGTCGNATATATGGGCTATTGGTTGTGTTTTTGCAGAGCTCCTGACAGGCCAG	594	
DB	541	TCCACAGCTGTGGACATCTGGAGCCCTGGGCTGCATCTTTTCTGAGATGTTGACTCGCCGG	600	
QY	595	CCACTGTGGCCTGGAATATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAG	652	
DB	601	GCCCTGTTCCCTGGAGATTCTGAGATTGACCAGCTCTTCGGATCTTTCCGGACICTGG	658	

RESULT 7

US-09-266-225D-13

; Sequence 13, Application US/09266225D

Patent No. 6573364

GENERAL INFORMATION:

APPLICANT: Nandabalan, Krishan

APPLICANT: Kingsmore, Stephen

; APPLICANT: Tchernev, Velizar

; TITLE OF INVENTION: Isolation and Characterization of Hermansky-Pudlak

TITLE OF INVENTION: Syndrome (HPS) Protein Complexes and HPS Protein-

; TITLE OF INVENTION: Interacting Proteins

; FILE REFERENCE: 15966-523

; CURRENT APPLICATION NUMBER: US/09/266,225D

; CURRENT FILING DATE: 1999-03-10

; NUMBER OF SEQ ID NOS: 19

```

; SOFTWARE: PatentIn Ver. 2.1

```

```

; SEQ ID NO 13
;
; LENGTH: 1476
;
; TYPE: DNA
;
; ORGANISM: Hmo sapiens
US-09-266-225D-13

```

Query Match: 15.6%; Score 162; DB 4; Length 1476;

Best Local Similarity 55.0%; Pred. No. 1, 1e-35;

Matches	362	Conservative	0	Mismatches	290	Indels	6	Gaps	2
---------	-----	--------------	---	------------	-----	--------	---	------	---

Qy	:	ATGGAAAAAGTATGAAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGGTGTATTCAAA	60
Db			
Qy	1	ATGGAGAACTTCCAAAAGGTGGAANAAGATCGGAGAGGGCA CGTACGGAGTGTGTACAAA	60
Db			
Qy	61	TGCAGAAAAACAAACCTCTGGACAAGTAGTAGCTGTAAAAAAATTTGTGGAATCTGAAGAT	120
Db			
Qy	61	GCCAGANAACAAGTTGACGGAGAGAGGTGGTGGCGCTTAAGAAAAATCCGCGCTGGACACTGAG	120
Db			
Qy	121	GATCCTGTTGTTAAGAAAAATAGCACTAAGAGAGAAATACGTATGTTGAAGCAATTAAAAACAT	180
Db			
Qy	121	ACTGAGGGTGTGCCCACTACTGCCATCCGAGAGATCTCTCTGCTTAAGGAGCTTAACCAT	180
Db			
Qy	181	CCAAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAAGGAAAATGCAATTAGTGTTTT	240
Db			
Qy	181	CCTAATATTGTCAAGCTGCTGGATGTCATTACACAGAAAAATAAACTCTACCTGGTGT	240
Db			
Qy	241	GAATACTGTGATCATACACTTTTAAATGAGCTGGA -- AAGAAACCCAAATGSAGTTGCT	297
Db			
Qy	241	GAATTTCTGCACCAAGATCTCAAGAAATTCATGATGCGCTCTGCTCTCACTGGCATTCCT	300
Db			
Qy	298	GATGGAGTGATCAAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAAATTTCTGTCAATATA	357
Db			
Qy	301	CTTCCCCTCATCAAGAGCTATCTGTCCAGCTGCTCCAGGGCCTAGCTTTCTGCCATTCT	360
Db			
Qy	358	CATAACTGTATTCACAGAGATATAAAACCTGAAAATATTCTAATACTAAGCAGGAATA	417
Db			
Qy	361	CATCGGGTCTCCACCAGACCTTAACCTCAGAACTGTCTTATTAACACAGAGGGGGCC	420
Db			
Qy	418	ATCAAGATTTGTGACTTCGGGTTTGACAAAATTCG -- ATTCCAGGAGATGCCCTACACC	474
Db			
Qy	421	ATCAAGCTAGCAGACITTTGGACTAGCCAGAGCTTTTGGAGTCCCTGTTCGTACTTACACC	480
Db			
Qy	475	GATTATGTAGCTACGAGATGGTACCGAGCTCCTGAACTTCTTGTGGGAGATACCTAGTAT	534
Db			
Qy	481	CATGAGGTGGTGACCCCTGTGGTACCGAGCTCCTGAAAATCCTCTCTGGGCTCGAATATTAT	540
Db			
Qy	535	GGTCTTTCAGTCGATATATGGGCTATTGGTGTGTTTTTTCAGAGAGCTCCTGCAGAGCCAG	594
Db			
Qy	541	TCCACAGCTGTGGACA TCTGGAGCCTGGGCTGCATCTTTGCTGAGATGGTGACTCGCCGG	600
Db			
Qy	595	CCACTGTGGCCTGGAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAG	652
Db			
Qy	601	GCCCTGTTCCCTGGAGATCTGAGATTGACCAAGCTCTTCCGGATCTTTCCGAGCTCTGG	658
Db			

RESULT 8

US-09-417-197-112

; Sequence 112, Application US/09417197

Patent No. 6518021

GENERAL INFORMATION:

APPLICANT: Ole THASTRUP, et al.

; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An

; TITLE OF INVENTION: On A Cellular Response

; FILE REFERENCE: 3759-0110P

; CURRENT APPLICATION NUMBER: US/09/417,197

; CURRENT FILING DATE: 1999-10-07

; NUMBER OF SEQ ID NOS: 143

```

; SOFTWARE: PatentIn version 3.0

```

; SEQ ID NO 112

; LENGTH: 1635

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: CDK2-EGFP fusion
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1632)
US-09-417-197-112

Query Match 15.6%; Score 162; DB 4; Length 1635;
Best Local Similarity 55.0%; Pred. No. 1.le-35;
Matches 362; Conservative 0; Mismatches 290; Indels 6; Gaps 2;

QY 1 ATGGAAAGTATGAAAAATTAGCTAAGACTGGAGAAGGGTCTTATGGGGTTGTATTCAAA 60
|||||
Db 1 ATGGAGAACTCCAAAAGGTGGAAAAGATCGGAGAGGGCACGTACGGAGTTGTGTACAAA 60
QY 61 TGCAGAAACAAAACCTCTGGACAAGTAGTAGCTGTTTAAAAAATTGTGGAATCTGAAGAT 120
|||||
Db 61 GCCAGAAACAAGTTGACGGGAGAGGTGGTGGCGCTTAAGAAAAATCCGCCCTGGACACTGAG 120
QY 121 GATCCTGTTGTTAAGAAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTTAAACAT 180
|||||
Db 121 ACTGAGGGTGTGCCCAGTACTGCCATCCGAGAGATCTCTCTGCTTAAGGAGCTTAACCAT 180
QY 181 CCAATCTTGTGAACCTCATCGAGGTGTTCAAGAGAAAAAGGAAAAATGCATTTAGTTTTT 240
|||||
Db 181 CCTAATATTGTCAAGCTGCTGGATGTCATTCACACAGAAAAATAAACTCTACCTGGTTTTT 240
QY 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGA---AAGAAACCCAAATGGAGTTGCT 297
|||||
Db 241 GAATTTCTGCACCAAGATCTCAAGAAATTCATGGATGCCCTCTGCTCTCACTGGCATTCCT 300
QY 298 GATGGAGTGATCAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGTGCATATA 357
|||||
Db 301 CTTCGCCCTCATCAAGAGCTATCTGTTCCAGCTGCTCCAGGGCTAGCTTTCTGCCATCT 360
QY 358 CATAACTGTATTACACAGAGATATAAAACCTGAAAAATATTCTAATACTAAGCAAGGAATA 417
|||||
Db 361 CATCGGGTCTCCACCGAGACCTTAAACCTCAGAAATCTGCTATTAAACACAGAGGGGGCC 420
QY 418 ATCAAGATTGTGACTTCGGGTTTGCACAAATTCGT---ATCCAGGAGATGCCCTACACC 474
|||||
Db 421 ATCAAGCTAGCAGACTTTGGACTAGCCAGAGCTTTTGGAGTCCCTGTTCTGACTTACACC 480
QY 475 GATTATGTAAGTACGAGATGGTACCGAGCTCCTGAACCTCTGTGGGAGATACCTCAGTAT 534
|||||
Db 481 CATGAGGTGTAACCTGTGGTACCGAGCTCCTGAATCCCTCTGGGCTCGAAATATTAT 540
QY 535 GGTCTTTCAGTCGATATATGGCTATTGGTGTGTGTTTTTGCAGAGCTCCTGACAGGCCAG 594
|||||
Db 541 TCCACAGCTGTGGACATCTGGAGCCTGGGCTGCATCTTTTGTGAGATGGTGAATCGCCGG 600
QY 595 CCACTGTGGCTGGAAAAATCAGATGTGGACCAACTTTTATCTGATAATCAGAAACACTAG 652
|||||
Db 601 GCCCTGTTCCCTGGAGATTCTGAGATTGACCAGCTCTTCCGGATCTTTTCGGACTCTGG 658

RESULT 9
US-09-417-197-114
; Sequence 114, Application US/09417197
; Patent No. 6518021
; GENERAL INFORMATION:
; APPLICANT: Ole THASTRUP, et al.
; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An
; FILE REFERENCE: 3759-0110P
; CURRENT APPLICATION NUMBER: US/09/417,197
; CURRENT FILING DATE: 1999-10-07
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 114
; LENGTH: 1635
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:

; OTHER INFORMATION: EGFP-CDK2 fusion
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1632)
US-09-417-197-114

Query Match 15.6%; Score 162; DB 4; Length 1635;
Best Local Similarity 55.0%; Pred. No. 1.le-35;
Matches 362; Conservative 0; Mismatches 290; Indels 6; Gaps 2;

QY 1 ATGGAAAGTATGAAAAATTAGCTAAGACTGGAGAAGGGTCTTATGGGGTTGTATTCAAA 60
|||||
Db 739 ATGGAGAACTTCCTAAAAGGTGGAAAAGATCGGAGAGGGCACGTACGGAGTTGTGTACAAA 798
QY 61 TGCAGAAACAAAACCTCTGGACAAGTAGTAGCTGTTTAAAAAATTGTGGAATCTGAAGAT 120
|||||
Db 799 GCCAGAAACAAGTTGACGGGAGAGGTGGTGGCGCTTAAGAAAAATCCGCCCTGGACACTGAG 858
QY 121 GATCCTGTTGTTAAGAAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTTAAACAT 180
|||||
Db 859 ACTGAGGGTGTGCCCAGTACTGCCATCCGAGAGATCTCTCTGCTTAAGGAGCTTAACCAT 918
QY 181 CCAATCTTGTGAACCTCATCGAGGTGTTCAAGAGAAAAAGGAAAAATGCATTTAGTTTTT 240
|||||
Db 919 CCTAATATTGTCAAGCTGCTGGATGTCATTCACACAGAAAAATAAACTCTACCTGGTTTTT 978
QY 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGA---AAGAAACCCAAATGGAGTTGCT 297
|||||
Db 979 GAATTTCTGCACCAAGATCTCAAGAAATTCATGGATGCCCTCTGCTCTCACTGGCATTCCT 1038
QY 298 GATGGAGTGATCAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGTGCATATA 357
|||||
Db 1039 CTTCGCCCTCATCAAGAGCTATCTGTTCCAGCTGCTCCAGGGCTAGCTTTCTGCCATCT 1098
QY 358 CATAACTGTATTACACAGAGATATAAAACCTGAAAAATATTCTAATACTAAGCAAGGAATA 417
|||||
Db 1099 CATCGGGTCTCCACCGAGACCTTAAACCTCAGAAATCTGCTATTAAACACAGAGGGGGCC 1158
QY 418 ATCAAGATTGTGACTTCGGGTTTGCACAAATTCGT---ATCCAGGAGATGCCCTACACC 474
|||||
Db 1159 ATCAAGCTAGCAGACTTTGGACTAGCCAGAGCTTTTGGAGTCCCTGTTCTGACTTACACC 1218
QY 475 GATTATGTAAGTACGAGATGGTACCGAGCTCCTGAACCTCTGTGGGAGATACCTCAGTAT 534
|||||
Db 1219 CATGAGGTGTAACCTGTGGTACCGAGCTCCTGAATCCCTCTGGGCTCGAAATATTAT 1278
QY 535 GGTCTTTCAGTCGATATATGGCTATTGGTGTGTGTTTTTGCAGAGCTCCTGACAGGCCAG 594
|||||
Db 1279 TCCACAGCTGTGGACATCTGGAGCCTGGGCTGCATCTTTTGTGAGATGGTGAATCGCCGG 1338
QY 595 CCACTGTGGCTGGAAAAATCAGATGTGGACCAACTTTTATCTGATAATCAGAAACACTAG 652
|||||
Db 1339 GCCCTGTTCCCTGGAGATTCTGAGATTGACCAGCTCTTCCGGATCTTTTCGGACTCTGG 1396

RESULT 10
US-09-220-132-3
; Sequence 3, Application US/09220132
; Patent No. 6506607
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT
; FILE REFERENCE: 07334-074001
; CURRENT APPLICATION NUMBER: US/09/220,132
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: US 60/079,303
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: US 60/068,821
; PRIOR FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3

DB 918 GCCCTGTTCTCTGGTGACTCTGAGATTGACCAGCTCTTTTCGTATCTTTTCGTATGCTGG 975

RESULT 12

US-08-154-915-1
; Sequence 1, Application US/08154915
; Patent No. 5618669
; GENERAL INFORMATION:
; APPLICANT: Beach, David
; APPLICANT: Xiong, Yue
; TITLE OF INVENTION: Cyclin Complex Rearrangement and Uses
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/154,915
FILING DATE: 19-NOV-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/991,997
FILING DATE: 17-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/963,308
FILING DATE: 16-OCT-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/888,178
FILING DATE: 26-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/702,514
FILING DATE: 16-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: MII-026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1089 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 13..888

US-08-154-915-1
Query Match 12.8%; Score 133.2; DB 1; Length 1089;
Best Local Similarity 52.3%; Pred. No. 11e-27;
Matches 344; Conservative 0; Mismatches 308; Indels 6; Gaps 2;
QY 1 ATGGAAAAGTATGAAAAAATTAGCTAAGACTGGAGAAGGTCCTTATGGGTTGTATTCAA 60
DB 13 ATGCAGAAATACGAGAAACTGGAAAAGATTGGGAAGGCACCTACGGAAGTGTGTTCAAG 72
QY 61 TGCAGAAACAAAACCTCTGGACAAGTAGTAGCTGTTAAAAAAATTTGTGAATCTGAAGAT 120
DB 73 GCCAAAAACCCGGGAGACTCATGAGATCGTGGCTCTAAAACCGGTGAGGCTGGATGACGAT 132
QY 121 GATCCTGTTGTTAAGAAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAATAACAT 180

DB 133 GATGAGGGTGTCCGAGTTCCGCCCTCCGGAGATCTGCCTACTCAAGGAGCTGAAGCAC 192
QY 181 CCAAATCTTGTGAACCTCATCGAGGTGTTCCAGGAGAAAAAGGAAAAATGCAATTTAGTTTTT 240
DB 193 AAGAACATCGTCAGGCTTCATGACGTCCTGCACAGCGACAAGAGCTGACTTTGGTTTTT 252
QY 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAGAAAGAAACCCAAATGGAGTTGCTGAT 300
DB 253 GAATTTCTGTGACCCAGGACCTGAAGAAGTATTTTGACAGTTGCAATGGTGACCTCGATCCT 312
QY 301 GGAGTGATCAAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGCATATACAT 360
DB 313 GAGATTGTAAGTCAATTCCTCTCTCCAGCTACTAAAAGGGCTGGGATTCTGTCTATAGCCGC 372
QY 361 AACTGTATTCACAGAGATATAAAACCTGAAAAATATTCTATAAATAAGCAAGGAATAATC 420
DB 373 AATGTGTACACAGGGACCTGAAGCCCCAGAACCTGCTATAAAACAGGAATGGGGAGCTG 432
QY 421 AAGATTTGTGACTTCGGGTTTGCACAAA---TTCTGATTCAGGAGATGCCTACACCGAT 477
DB 433 AAATGGCTGATTTTGGCCTGGCTCGAGCCCTTGGGATTCCTCGCTCGCTGTTACTCAGCT 492
QY 478 TATGTAGCTACGAGATGGTACCAGACTCCTGAACCTTCTGTGGGAGATACTCAGTATGGT 537
DB 493 GAGTGTGCACACTGTGTGTACCGCCACCGGATGTCTCTTTGGGGCCAAAGCTGTACTCC 552
QY 538 TCTTCAGTCTGATATATGGGCTATTGGTTGTGTTTTTGCAGAGCT---CCTGACAGGCCAG 594
DB 553 ACGTCCATCGACATGTGTGACGCCGCTGCATCTTTGCAGAGCTGGCCAATGCTGGCGG 612
QY 595 CCACTGTGGCCTGGAATAATCAGATGTGGACCAACTTTTATCTGATAAATCAGAACACTAG 652
DB 613 CTTCTTTTCCCGCAATGATGTCGATGACCAGTGAAGAGGATCTTCCGACTGCTGG 670

RESULT 13

US-08-464-517-37
; Sequence 37, Application US/08464517
; Patent No. 5869640
; GENERAL INFORMATION:
; APPLICANT: BEACH, David H.
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,517
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/963,308
; FILING DATE: 16-OCT-1992
; APPLICATION NUMBER: US 07/888,178
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/701,514
; FILING DATE: 16-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Matthew P. Vincent
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-004C
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 1089 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 13..888
US-08-464-517-37

Query Match 12.8%; Score 133.2; DB 2; Length 1089;
Best Local Similarity 52.3%; Pred. No. 1.1e-27;
Matches 344; Conservative 0; Mismatches 308; Indels 6; Gaps 2;

QY 1 ATGGAAGATGATAAATAGCTAAGACTGGAGAAGGGCTTATGGGTTGTATTCAA 60
DB 13 ATGCAGAAATACGAGAAACTGGAAAGATTTGGGAAGGCCCTACGGAACGTGTGTTCAAG 72

QY 61 TGCAGAAACAAACCTCTGGACAGTAGTAGCTGTTAAATAATTTGTGGAATCTGAAGAT 120
DB 73 GCCAAACACGGGAGACTCATGATCGTGGCTCTAAACCGGTGAGGCTGGATGACGAT 132

QY 121 GATCCTGTGTTAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAAT 180
DB 133 GATGAGGTGTCCGAGTTCAGCTGCTCCGGGAGATCTGCCTACTCAAGGAGCTGAAGCAC 192

QY 181 CCAATCTGTGAACCTCATCGAGGTGTTCAAGGAGAAAGGAAATAGCATTTAGTTT 240
DB 193 AAGAACATCGTCAGGCTTCATGACGCTCTGCACAGCGACAAGAGTGAATGTTGTTT 252

QY 241 GAATACCTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGGAGTTGCTGAT 300
DB 253 GAATTCGTGACCAAGGACCTGAAGAGTATTTTGACAGTTGCAATGGTGACCTCGATCCT 312

QY 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAGCTCTTAATTTCTGTCATATACAT 360
DB 313 GAGATTGTAAAGTCATTCCTCTCCAGCTACTAAAGGGCTGGGATCTGTGCATAGCCGC 372

QY 361 AACTGTATTCACAGAGATATAAACCTGAAATATTTCTAATAACTAAGCAAGGAATAATC 420
DB 373 AATGTGCTACACAGGGACCTGAAGGCCCAACCTGCTAATAACAGGAATGGGAGCTG 432

QY 421 AAGATTTGTGACTTCGGGTTTGACAAA---TTCTGATTCAGAGATGCCTACACCGAT 477
DB 433 AAATTGGCTGATTTTGGCTGGCTCGAGGCTTTGGGATTCCTGCTCCGCTGCTGCTGCT 492

QY 478 TATGTAGCTACAGAGATGGTACCGAGCTCCTGAACCTCTTGCGAGATACCTCAGTATGGT 537
DB 493 GAGGTGGTCACACTGTGGTACCGCCCGGCTGCATCTTTGCAGAGCTGGCCAAATGCTGGCGG 552

QY 538 TCTCAGTCGATATATGGGCTATTGGTGTGTTTTCAGAGCT---CCTGACAGGCCAG 594
DB 553 ACGTCCATCGACATGTGGTCAGCCGGCTGCATCTTTGCAGAGCTGGCCAAATGCTGGCGG 612

QY 595 CCACTGTGGCCTGGAAATCAGATGTGGACCAACTTTATCTGTATATCAGAACACTAG 652
DB 613 CCTCTTTTCCCGCAATGATGTCGATGACCAAGTTGAAGAGGATCTTCGACACTGCTGG 670

RESULT 14
US-08-246-361A-37
Sequence 37, Application US/08246361A
Patent No. 5998582
GENERAL INFORMATION:
APPLICANT: BEACH, David H.
TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD

STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/246,361A
FILING DATE: 19-MAY-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/963,308
FILING DATE: 16-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/888,178
FILING DATE: 26-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/701,514
FILING DATE: 16-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Matthew P. Vincent
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: MII-004C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 1089 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 13..888
US-08-246-361A-37

Query Match 12.8%; Score 133.2; DB 2; Length 1089;
Best Local Similarity 52.3%; Pred. No. 1.1e-27;
Matches 344; Conservative 0; Mismatches 308; Indels 6; Gaps 2;

QY 1 ATGGAAGATGATAAATAGCTAAGACTGGAGAAGGGCTTATGGGTTGTATTCAA 60
DB 13 ATGCAGAAATACGAGAAACTGGAAAGATTTGGGAAGGCCCTACGGAACGTGTGTTCAAG 72

QY 61 TGCAGAAACAAACCTCTGGACAAAGTAGTAGCTGTTAAATAATTTGTGGAATCTGAAGAT 120
DB 73 GCCAAACACGGGAGACTCATGAGATCGTGGCTCTAAACCGGTGAGGCTGGATGACGAT 132

QY 121 GATCCTGTGTTAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAAT 180
DB 133 GATGAGGTGTCCGAGTTCAGCTGCTCCGGGAGATCTGCCTACTCAAGGAGCTGAAGCAC 192

QY 181 CCAATCTGTGAACCTCATCGAGGTGTTCAAGGAGAAAGGAAATAGCATTTAGTTT 240
DB 193 AAGAACATCGTCAGGCTTCATGACGCTCTGCACAGCGACAAGAGCTGACTTTGGTTTT 252

QY 241 GAATACCTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGGAGTTGCTGAT 300
DB 253 GAATTCGTGACCAAGGACCTGAAGAGTATTTTGACAGTTGCAATGGTGACCTCGATCCT 312

QY 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAGCTCTTAATTTCTGTCATATACAT 360
DB 313 GAGATTGTAAAGTCATTCCTCTTCAGCTACTAAAGGGCTGGGATTTCTGTCATAGCCGC 372

QY 361 AACTGTATTCACAGAGATATAAACCTGAAATATTTCTAATAACTAAGCAAGGAATAATC 420
DB 373 AATGTGCTACACAGGGACCTGAAGGCCCAACCTGCTAATAACAGGAATGGGAGCTG 432

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2003, 16:06:46 ; Search time 2620.85 Seconds
(without alignments)
9653.738 Million cell updates/sec

Title: US-09-671-050-9

Perfect score: 1041

Sequence: 1 atggaaaagtatgaaaaatt.....aggtacttcgctcaaaagt 1041

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estmu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_htc:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_htc:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: em_gss_hum:*
- 18: em_gss_inv:*
- 19: em_gss_pln:*
- 20: em_gss_vrt:*
- 21: em_gss_fun:*
- 22: em_gss_mam:*
- 23: em_gss_mus:*
- 24: em_gss_pro:*
- 25: em_gss_rtd:*
- 26: em_gss_phg:*
- 27: em_gss_vr1:*
- 28: gb_gss1:*
- 29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
C 1	480.6	46.2	695	14	CB169554	CB169554 RUC603000
2	405	38.9	670	14	BY733578	BY733578 BY733578
3	360.2	34.6	689	14	BY752739	BY752739 BY752739
4	333.6	32.0	500	9	AI385966	AI385966 ml34h09.y

5	331	31.8	1691	11	AK016781	AK016781 Mus muscu
6	316.8	30.4	938	9	AA061797	AA061797 ml34h09.y
7	306.8	29.5	598	9	AV986182	AV986182 AV986182
8	298.4	28.7	824	13	BU221831	BU221831 603750354
C 9	295.6	28.4	536	13	BU686325	BU686325 UI-CF-DU1
10	289.2	27.8	597	12	BP016186	BP016186 BP016186
11	284.6	27.3	825	14	CB315094	CB315094 AGENCOURT
12	269.8	25.9	565	14	CB400506	CB400506 OSTFI77F1
13	258.6	24.8	731	9	AW106692	AW106692 um32d03.y
14	255	24.5	718	9	AA286088	AA286088 vc33b05.y
15	249	23.9	508	12	B1511582	B1511582 BB160006A
16	247.6	23.8	512	13	BX304025	BX304025 BX304025
17	239	23.0	485	9	AV960213	AV960213 AV960213
18	223.8	21.5	688	13	BW241505	BW241505 BW241505
19	221.8	21.3	2572	11	AK030598	AK030598 Mus muscu
20	213.4	20.5	570	12	BJ000534	BJ000534 BJ000534
21	213	20.5	1064	13	BU120363	BU120363 603142234
22	210.8	20.2	627	14	BY716459	BY716459 BY716459
23	209.8	20.2	579	9	AW233105	AW233105 fj28b10.y
24	209.4	20.1	380	13	BY075395	BY075395 BY075395
25	202.2	19.4	2005	11	AK090045	AK090045 Mus muscu
26	202.2	19.4	2769	11	AK045036	AK045036 Mus muscu
27	200.6	19.3	1893	11	AK050990	AK050990 Mus muscu
C 28	199	19.1	295	9	AA626859	AA626859 zu89f10.s
29	199	19.1	2388	11	AK046394	AK046394 Mus muscu
30	199	19.1	2665	11	AK045356	AK045356 Mus muscu
31	197.2	18.9	471	14	CB732963	CB732963 AMGNNUC:N
32	189.6	18.2	2013	11	AK078804	AK078804 Mus muscu
33	188.6	18.1	659	13	BW280090	BW280090 BW280090
34	187	18.0	673	13	BW036816	BW036816 BW036816
35	186.6	17.9	727	9	AI508835	AI508835 vc33b05.y
36	186.2	17.9	666	14	BY733092	BY733092 BY733092
37	185.2	17.8	2604	11	AK052380	AK052380 Mus muscu
C 38	183.4	17.6	645	14	CA334835	CA334835 NISC_lt01
39	183.4	17.6	812	13	BU377342	BU377342 603812152
40	182	17.5	320	13	BY121897	BY121897 BY121897
C 41	180.6	17.3	774	13	BW110560	BW110560 BW110560
42	175.6	16.9	1535	11	BC010966	BC010966 Homo sapi
43	174.8	16.8	363	10	BB868874	BB868874 BB868874
44	174.8	16.8	582	9	AV672492	AV672492 AV672492
45	174.4	16.8	665	28	AQ953706	AQ953706 Sheared D

ALIGNMENTS

RESULT : CB169554/c
LOCUS CB169554 695 bp mRNA linear EST 30-JAN-2003
DEFINITION RUC603000752.R1 CSEQFXN20 testes Bos taurus cDNA, mRNA sequence.
ACCESSION CB169554
VERSION CB169554.1 GI:28155681
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus

REFERENCE 1 (bases 1 to 695)
AUTHORS Adelson,D.L. and Gill,C.A.
TITLE Bovine ESTs (Adelson and Gill)
JOURNAL Unpublished
COMMENT Contact: David L. Adelson
Animal Breeding and Genetics
Texas A&M University
Animal Science Dept., TAMU-2471, College Station, TX 77843-2471,
USA
Tel: 9798452616
Fax: 9798456970
Email: david.adelson@tamu.edu.
Location/Qualifiers
1..695
/organism="Bos taurus"

FEATURES
Source


```
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="testes"
/clone_fib="CSEQFXN20 testes"
/note="Organ: testes; Vector: pBluescript SK+; Site_1:
Not1: Site_2: EcoRI; sequence 5' of the insert
(5'-NNN...NNNinsert)
CGCAATTGGAGCTCCACCGCGGTGGCGCGCGGCTCGAG. Sequence 3' of
the inserts (AAGATTCGATATCAAGCTATCGATACCGCTCGACCTCGAG.
normalized Rd 1 library, sequenced 3' with M13R primer."
BASE COUNT      169 a   166 c   133 g   227 t
ORIGIN
Query Match      46.2%; Score 480.6; DB 14; Length 695;
Best Local Similarity 92.8%; Pred. No. 3.3e-22;
Matches 504; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
QY 1 ATGGAAGAAGTATGAAGAATTAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAA 60
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 543 ATGGAGAAGTATGAAGAATTAGCTAAGATTGGGGAGGGTCTTATGGGTTGTATTCAA 484
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 TGCAGAACAAACCTCTGGACAAGTAGTAGCTGTTTAAAAAATTTCGTGGATCTGAAGAT 128
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 483 TGCAGAACAAACCTCGGACAAGTGTAGCTATTAAAAAATTTCGTGGATCCGAAGAT 424
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 121 GATCCTGTGTTAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAT 180
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 423 GATCCTGTGTTAAGAAATAGCTCTAAGAGAAATACCATGTTAAAGCAATTAACAT 364
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 181 CCAATCTGTGAACCTCATCGAGGTGTTTCAGGAGAAAGAAATGCATTTAGTTTTT 240
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 363 CCAATCTGTGAACCTCATCGAGGTGTTTCAGGAGAAAGAAATGCATTTAGTTTTT 304
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 241 GAATACCTGTGATCATACACTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT 300
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 303 GAGTACTGTGATCACACACTTTTAAATGAGCTGGAAAGAAACCCAAACGGAGTTGCTGAT 244
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 301 GGAGTGATCAAAAGCGTATTATGCGCAACACTTCAAGCTCTTAATTTCTGTATATACAT 360
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 243 GGAGTGATCAAAAGTGTGTTATGGCAACACTCCAAAGCTCTCAACTTCTGTATATAACAC 184
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 361 AACTGTATTACAGAGATATAAAACCTGAAATATTCTAATAACTAAGCAAGGAATAATC 420
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 183 AATTGTATTACAGAGATGTAAACCTGAAATATTCTAATAACTAAGCAAGGAATAATC 124
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 421 AAGATTTGTGACTTCGGGTTTGCAAAATTTCTGATTCAGGAGATGCCTACACCGATTAT 480
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 123 AAGATCTGTGACTTTGGGTTTGCAAGAAATTTCTGATTCAGGAGATGCCTACACTGACTAC 64
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 481 GTAGCTACGAGATGGTACCGAGCTCCTGAACCTTCTTGTGGAGATACCTCAGTATGTTCT 540
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 63 GTCGCTACAAGATGGTACCGAGCTCCCGACCTTCTTGTGGGGGATACCCAGTATGTTCT 4
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 541 TCA 543
Db |||
QY 3 TCA 1
```

```
RESULT 2
BY733578
LOCUS
DEFINITION
BY733578 RIKEN full-length enriched, 16 days neonate male
diencephalon Mus musculus cDNA clone G630052E12 5', mRNA sequence.
ACCESSION
BY733578
VERSION
BY733578
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 670)
AUTHORS
Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I., Kiyosawa,H.
```

Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A., Schonbach,C.,
Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C., Hume,D.A.,
Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H., Batalov,S.,
Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V., Chothia,C., Corbani
L.E., Cousins,S., Dalla,E., Dragani,T.A., Fletcher,C.F., Forrest
A., Frazer,K.S., Gaasterland,T., Gariboldi,M., Gissi,C., Godzik,A.,
Gough,J., Grimmond,S., Gustincich,S., Hirokawa,N., Jackson,I.J.,
Jarvis,E.D., Kanai,A., Kawaji,H., Kawasawa,Y., Kedzierski,R.M.,
King,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons
P.A., Maglott,D.R., Maltais,L., Marchionni,L., McKenzie,L., Miki
H., Nagashima,T., Numata,K., Okido,T., Pavan,W.J., Pertea,G.,
Pesole,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D.,
Ramachandran,S., Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring
B.Z., Ringwald,M., Sandelin,A., Schneider,C., Semple,C.A., Setou
X., Shimada,K., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale
R.D., Tomita,M., Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y.,
Watanabe,Y., Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa
M., Yang,I., Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A.,
Carninci,P., Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura
M., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii
Y., Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata
K., Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander
E.S., Rogers,J., Sirney,E. and Hayashizaki,Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
22354683
12466851
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploitation Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda
S., Hashizume,M., Hayashida,K., Hirozane,T., Hori,F., Imotani,K.,
Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y., Kondo,S., Konno
H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M., Nomura,K.,
Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Sakazume,N., Sano,H.,
Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M., Takeda,Y.,
Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y. Direct
Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.
Location/Qualifiers
i. .670
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="G630052E12"
/sex="male"
/tissue_type="diencephalon"

FEATURES
source

```
/dev_stage="16 days neonate"
/clone_lib="RIKEN full-length enriched, 16 days neonate
male diencephalon"
BASE COUNT      205 a      140 c      169 g      155 t      1 others
ORIGIN

Query Match      38.9%;   Score 405;   DB 14;   Length 670;
Best Local Similarity 85.6%;   Pred. No. 1.8e-17;
Matches 450;   Conservative 0;   Mismatches 76;   Indels 0;   Gaps 0;

QY 1 ATGGAAAGTATGAAATAATAGCTAAGACTGAGAGAGGCTTTATGGGTTGTATTCAAA 60
    |||||
Db 145 ATGGAAAGTATGAAAGCTAGCTAAGATCGGAGAGGGCTTTATGGCGTTGTATTCAG 204

QY 61 TGCAGAAACAAACCTCTGGACAGTAGTAGCTGTTAAATAATTTGTGGATCTGAAGAT 120
    |||||
Db 205 TGCAGAAACAAATCTTCTGGACAGTAGTAGCGATCAAAAATTCGTGGAATCTGAAGAT 264

QY 121 GATCCTGTGTTAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAT 180
    |||||
Db 265 GATCGTGTGTTAGGAAATAGCCCTCGGGAAATCCGTATGCTGAAGCAGTTGAACAC 324

QY 181 CCAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAGAGAGAGAGAGAGAGAGAG 240
    |||||
Db 325 CCAACCTCTGTGAACCTCATCGAGGTGTTTCAGAGAGAGAGAGAGAGAGAGAGAG 384

QY 241 GAATCTGTGATCATACACTTTTAAATGAGCTGGAAGAGAAACCCAAATGGAGTTGCTGAT 300
    |||||
Db 385 GAGTACTGTGATCACACACTGTTAAACGAGCTGGAGAGAGAGAGAGAGAGAGAGAG 444

QY 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAGCTTCAATTTCTGTCTATATACAT 360
    |||||
Db 445 GGAGTGATTAAGAGTGTGCTATGGCAACCCCTTCAAGCCCTTAACTTCTGTCAAGAC 504

QY 361 AACTGTATTCACAGAGATATAAAACCTGAAATAATTTCTAATACTAAGCAAGGAATAATC 420
    |||||
Db 505 AATTGTATTCATCGGGATGTAACCTGAAACATCTCTAATAACNAGCAAGGGATGATA 564

QY 421 AGATTGTTGACTTCGGGTTTGACAAATTCCTGATTCAGGAGATGCCTACACCGATTAT 480
    |||||
Db 565 AAGATTGTTGACTTTGGATTGTCAGGAATCTTAATTCAGGAGAGCGCTACACAGACTAT 624

QY 481 GTAGCTACGAGATGTTACCGAGCTCTGAACTTCTTGTGGGAGATA 526
    |||||
Db 625 GTTCCACCAGTGGTACCGAGCCCNCGAACTTCTCTGTGGAGACA 670

RESULT 3
BY752739
LOCUS
DEFINITION
BY752739 RIKEN full-length enriched, adult inner ear Mus musculus
CDNA clone F930014106 5', mRNA sequence.
ACCESSION
BY752739
VERSION
BY752739.1 GI:27183802
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 689)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H.,
Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C.,
Gojchori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A.,
Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S.,
Beisel, K.W., Blake, J.A., Bradt, D., Brusio, V., Chothia, C., Corbani
, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest
, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,
Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J.,
Jarvis, E.D., Kanai, A., Kawaji, H., Kawasaki, Y., Kedzierski, R.M.,
King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons
, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki
, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G.,
```

```
Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D.,
Ravachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring
, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Sempke, C.A., Setou
, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale
, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y.,
Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa
, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A.,
Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura
, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
, Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii
, Y., Itch, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata
, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander
, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
22354683
12466851
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda
, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F., Imotani, K.,
Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y., Kondo, S., Konno
, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K.,
Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N., Sano, H.,
Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y.,
Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct
Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Kirk W. Beisel ( Boys Town National
Research Hospital 555 North 30th Street Omaha, NE 68131 USA ) whose
assistance we gratefully acknowledge.
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
FEATURES
Location/Qualifiers
1..689
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="CS7BL/6J"
/db_xref="taxon:10090"
/clone="F930014106"
/tissue_type="inner ear"
/dev_stage="adult"
/clone_lib="RIKEN full-length enriched, adult inner ear"
BASE COUNT      213 a      133 c      171 g      168 t      4 others
ORIGIN
Query Match      34.6%;   Score 360.2;   DB 14;   Length 689;
Best Local Similarity 82.7%;   Pred. No. 1.1e-14;
Matches 421;   Conservative 0;   Mismatches 87;   Indels 1;   Gaps 1;
```

QY 1 ATGGAAGAAGTATGAAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGGTTGTATTCAA 60
Db 169 ATGGAAGAAGTATGAAAAAGCTAGCTAAGATCGGAGAGGGTCTTATGGCGTTGTATTCAA 228

QY 61 TGCAGAAACAAACCTCTGGACAAGTAGTAGCTGTATAAATAATTTGTGGAATCTGAAGAT 120
Db 229 TGCAGAAACAAATCTTTCGGACAAGTAGTAGCGATCAAAAAATTCGTGGAATCTGAAGAT 288

QY 121 GATCCTGTGTTAAGAAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAACAT 180
Db 289 GATCGTGTCTTAGGAAAAATAGCCCTCGCGGAAATCCGTATGCTGAAGCAGTTGAACAC 348

QY 181 CAAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAAGGAAATGCAATTTAGTTT 240
Db 349 CCAACCTCGTGAACCTCATTCAGGTGTTTCAGAGAAAGAGGAAGATGCATCTAGTTT 408

QY 241 GAATACCTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGGAGTTGCTGAT 300
Db 409 GAGTACTGTGATCACACGCTGTTAAACGAGCTGGAGAGAAACCCAAACGGAGTTTCTGAT 468

QY 301 GGAGTGATCAAAAGCGTATATATGGCAAAACACTTCAAGCTCTTAATTTCTGTCAATACAT 360
Db 469 GGAGTGATTAAGAAGTGTGCTATGGCAAAACCCCTTCAAGCCCTTAACCTTCTGTCAAGCAC 528

QY 361 AACTGTATTCACAGAGATATAAAACCTGAAATATATTTCTAATAAGCAAGGAATAATC 420
Db 529 AATTGTATTCATCGGAGTGTAAACCTGAAACATCTTAATACCAAGCAGGGGATGANT 588

QY 421 AAGATTGTGACTTCGGGTTTGACAAATCTGTATCCAGGAGATGCTTACACCGGATTAT 480
Db 589 AAGATTGTGACTTTGGATTGTCACGATTTCTTAATNCA-GAGACGGCTTACACANGACTA 647

QY 461 GTAGCTACGAGATGGTACCGAGCTCCTGA 509
Db 648 TGTCCANCCAGTGGTACCGAGCCCCCGA 676

RESULT 4
AI385966 500 bp mRNA linear EST 27-JAN-1999
LOCUS ml34h09.y1 Stratagene mouse testis (#937308) Mus musculus cDNA
DEFINITION clone IMAGE:513953 5' similar to gb:X66358 SERINE/THREONINE-PROTEIN
KINASE KIAALRE (HUMAN);, mRNA sequence.
ACCESSION AI385966
VERSION AI385966.1 GI:4199429
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 500)
AUTHORS Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person ,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter ,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE The WashU-NCI Mouse EST Project 1999
JOURNAL Unpublished
COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through JLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:307801
This read is a RESEQUENCE of a previously sequenced mouse clone
This read has been verified (found to hit its original self in the correct orientation)
Seq primer: -40RP from Gibco
High quality sequence stop: 339.
Location/Qualifiers

source i. 500
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CJ-1"
/dc_xref="taxon:10090"
/clone="IMAGE:513953"
/sex="males"
/tissue_type="testis"
/dev_stage="10-12 week old"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="Stratagene mouse testis (#937308)"
/note="Organ: testis; Vector: pBluescript SK-; Site:1: EORI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAAATCGCACGAG 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTT 3'."
BASE COUNT 145 a 114 c 121 g 120 t
ORIGIN

Query Match 32.0%; Score 333.6; DB 9; Length 500;
Best Local Similarity 82.3%; Pred. No. 6.2e-13;
Matches 408; Conservative 0; Mismatches 84; Indels 4; Gaps 2;

QY 135 GAAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAACATCCAAATCTGTGAA 194
Db 2 GAAAAATAGCCCTCGCGGAAATCCGTATGCTGAAG---TTGAACAACCCAAACCTCGTGA 58

QY 195 CCTCATCGAGGTGTTTCAGGAGAAAAAGGAAATGCAATTTAGTTTGTGAATACCTGTGATCA 254
Db 59 CCTCATCGAGGTGTTTCAGAGAAAGAGAAAGATGCATCTAGTTTGTGAGTACTGTGATCA 118

QY 255 TACACTTTTAATGAGCTGGAGAAACCCAAATGGAGTTGCTGATGGAGTGATCAAAAG 314
Db 119 CACACTGTTAAACGAGCTGGAGAGAAACCCAAACCGAGTTTCTGATGGAGTGATTAAGAG 178

QY 315 CGTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGTCATATACATAAATGTATTACAG 374
Db 179 TGTGCTATGGCAAAACCCCTTCAAGCCCTTAACCTTCTGTCACAAGCACAAATTGTATTATCG 238

QY 375 AGATATAAAACCTGAAAAATATTCTAATAAATAAGCAAGGAATAANTCAAGATTGTGACTT 434
Db 239 GGATGTAACCTGAAAAACATCTCTAATAAACCAAGCAAGGATGATAAAGATTGTGACTT 298

QY 435 CGGGTTTGCAAAATCTGATTCAGGAGATGCCTACACCGATTATGTAGCTACGAGATG 494
Db 299 TGGATTTCACGAATCTTAATTCAGGAGACGCGCTACACAGACTATGTTGCCACCGAGTG 358

QY 495 GTACCGAGCTCCTGAACCTCTTGTGGGAGATACCTCAGTATGTTTCTTCAGTCGATATATG 554
Db 359 GTACCGAGCCCTCGAACTTCTCTGGGAGACACGAAGTACGGTTCTCTGTAGAGCTGTG 418

QY 555 GGCTATTGGTTGTGTTTTTTCAGAGCTCCTGACAGGCCAGCCACTGTGGCCCTGSAATAATC 614
Db 419 GGCCGTCGGCTGTGTTTATGACAGAGCTTCTGACGTTGTACGCCACTCTGGCCCGG-AAATA 477

QY 615 AGATGTGGACCAACTT 630
Db 478 CGACGTGGACCACTT 493

RESULT 5
AKC16781 1691 bp mRNA linear HTC 05-DEC-2002
LOCUS Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4933411017 product:cyclin-dependent kinase-like 1 (CDC2-related kinase), full insert sequence.
DEFINITION AKC16781
ACCESSION AK016781
VERSION AK016781.1 GI:12855701
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

Carninci,P. and Hayashizaki,Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

99279253

20349636

Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,

Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

20499374

11042159

Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,

Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,

Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,

Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,

Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,

Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,

Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer

Genome Res. 10 (11), 1757-1771 (2000)

20530913

11076861

Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,

Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,

Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,

Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,

Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,

Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,

Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,T., Pesole,G.,

Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,

Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,

Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,

Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,

Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,

Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,

Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,

Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,

Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,

Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,

Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawai,H., Kohtsuki,S.

and Hayashizaki,Y.

Functional annotation of a full-length mouse cDNA collection

Nature 409 (6821), 685-690 (2001)

21085660

11217851

The FANTOM Consortium and the RIKEN Genome Exploration Research

Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation

of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

6 (bases 1 to 1691)

Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H.,

Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y.,

Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K.,

Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M.,

Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,

Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K.,

Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C.,

Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,

Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y.,

Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T.,

Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K.,

Yoshino,M., Muramatsu,M. and Hayashizaki,Y.

Direct Submission

Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of

Physical and Chemical Research (RIKEN), Laboratory for Genome

COMMENT

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in RIKEN

Division of Experimental Animal Research in Riken contributed to

prepare mouse tissues. First strand cDNA was primed with a primer

[5' GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTTNN 3'], cDNA was

prepared by using trehalose thermo-activated reverse transcriptase

and subsequently enriched for full-length by cap-trapper. Second

strand cDNA was prepared with the primer adapter of sequence [5'

GAGAGAGAGATTCGAGTTAATTAAATTAATCCCCCCCCC 3']. cDNA was cleaved

with BamHI and XhoI. cDNA of size comprised longer than 7 kb was

selected before cloning. Vector: a modified pBluescript KS(+) after

bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3'

end: BamHI. Host: DH10B.

Location/Qualifiers

1..1691

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM DB:4933411017"

/db_xref="MGI:1903578"

/db_xref="taxon:10090"

/clone="4933411017"

/sex="male"

/tissue_type="testis"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="adult"

177..1235

/note="unnamed protein product; cyclin-dependent

kinase-like 1 (CDC2-related kinase) (MGD|MGI:1918341)

putative"

/codon_start=1

/protein_id="BAC25497.1"

/db_xref="GI:26382418"

/db_xref="MGI:1918341"

/translation="MEKYEKIGKIGESYGVVFKCRNPDGTQIVAIKRFLETFEDDPVI

KKIALREIRMLKQLKHPNLVNLLEVFRRKRRLHLVFEYCDHTLVHELDRYQGVPEPL

VKNITWQLQAVNFCCHKNCIHRDVKNILITKQSAIKLCDFGFARLLTGPQYYTD

YVATRWYSPPELLVGDYQGPVYDVAIGVFAELSSVPLWPVKSDVQLYLRKTL

GDLIPIRHOQVFSMNQYFSGVKIPDPEDMETLELKFPNISYSALGFLKGCLHMDPAERL

TCEQLQHPYFDSIREVGEITRQHDKPARKTLRQSRKHLTGLQYLPQLTSSRIPLALD

NKKYHCSTKRFNYHFPNI"

BASE COUNT 485 a 395 c 417 g 394 t

ORIGIN

Query Match 31.8%; Score 331; DB 11; Length 1691;

Best Local Similarity 70.0%; Pred. No. 4.7e-13;

Matches 461; Conservative 0; Mismatches 195; Indels 3; Gaps 1;

CY 1 ATGGAAAAGTATGAAAATAGCTAAGACTGGAGAAGGGTCTTATGGGGTTGATTCAA 60

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

DT 177 ATGGAAAATATGAAAATAATTGGAAAGATTGGAGAAGGCTCTTATGGGGTAGTGTTC 236

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

CY 61 TGCAGAAACAAACCTCTTGGACAAGTAGTAGCTGTTTAAAAAATTTGTGGAATCTGAAGAT 120

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

DB 237 TSCAGAAACAGGGACACGGGTGAGATCGTGGCCATCAAGAGGTTTCTGGAACCGAAGAT 296

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

CY 121 GATCCTGTTGTTAAGAAAATAGCACTAAGAGAAAATACGTAFTGTTGAAGCAATTAACAAT 180

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

DB 297 GACCTGTGTCATAAAGAAAATCGCCCTTCGAGAAATCCGCATGCTCAAGCAACTCAAGCAC 356

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

CY 181 CCAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAGGAAAATGCAATTAGTTTTT 240

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

DB 357 CCCAACCTGGTCAACCTCCTGGAAGTCTTCGGAGGAGCGGAGGCTTCACCTGGTGTTC 416

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

CY 241 GAATACTGTGATCATACACTTTTAAATGACCTGGAAAGAAACCCAAATGGAGTTGCTGAT 300

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 417 GAGTCTGCGACACACGGTGTTCACGAGCTGGATCGGTATCAGAGGGGTACCAGAG 476
QY 301 GGAGTGATCAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAATTTCTTCATATACAT 360
Db 477 CCTCTCGTGAAGAACATAACTTGGCAGACACTGCGAGCTGTAAATTTCTGCCATAAACAT 536
QY 362 AACTGTATTACAGAGATATAAAACCTGAAATATTCTAATAACTAAGCAAGGAATATC 420
Db 537 AACTGCATACAGAGACGCTGAACCGGAAATATTCTCATCACCACAGTCAGCCATT 596
QY 421 AAGATTGTGACTTCGGGTTTGGACAAATTCGTATT---CCAGGAGATGCTTACACCGAT 477
Db 597 AAGCTCTGTGACTTTGGGTTTCCGACGGCTTCTCACTGGACCGAGTGAAGTACTACACGGAC 655
QY 478 TATGTAGTACGAGATGGTACCGAGCTCCTGAACCTTCTTGTGGGAGATACCTAGTATGGT 537
Db 657 TACGTGGCCACCGGTTGGTACCGCTCACCCGAGCTGCTAGTGGGAGACAGCAGTATGGT 716
QY 538 TCTTCAGTGGATATATGGGCTATTGGTGTGTTTTTGCAGAGCTCCTTGACAGGCCAGCCA 597
Db 717 CCCCCTGTAGATGTCTGGGCAATTGGCTGTGTGTTGCTGAGTTACTGTGCGGAGTGCCT 776
QY 598 CTGTGGCTGGAAATCAGATGTGGACCAACTTTATCTGATATATCAGAACACTAGTAGA 655
Db 777 CTATGGCCAGGAAATCTGACGTGGATCAGCTCTACCTGATAGGAAACCCCTGGGGGA 835

RESULT 6
AA061797
LOCUS ml34h09.r1 Stratagene mouse testis (#937308) Mus musculus cDNA
DEFINITION clone IMAGE:513953 5' similar to GB:X66358 SERINE/THREONINE-PROTEIN KINASE KKIALRE (HUMAN);, mRNA sequence.

ACCESSION AA061797.1 GI:1555606
VERSION AA061797.1
KEYWORDS Mus musculus (house mouse)
SOURCE EST.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS 1 (bases 1 to 938)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:307801
Seq primer: -28ml3 rev1 ET from Amersham
High quality sequence stop: 442.

FEATURES
source
1..938
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CD-1"
/db_xref="taxon:10090"
/clone="IMAGE:513953"
/sex="males"
/tissue_type="testis"
/dev_stage="10-12 week old"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="Stratagene mouse testis (#937308)"

/note="Organ: testis; Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Average insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCAGAG 3' ~3' adaptor sequence: 5' CTCGAGTCTTTTCTTTTCTTTT 3'."
BASE COUNT 254 a 215 c 236 g 233 t
ORIGIN
Query Match 30.4%; Score 316.8; DB 9; Length 938;
Best Local Similarity 82.9%; Pred. No. 5e-12;
Matches 398; Conservative 0; Mismatches 77; Indels 5; Gaps 3;
QY 125 GAAAATAGCACTAAGAGAAATACGATGTTGAAGCAATTAAACATCCAAATCTTGTA 194
Db 1 GAAAATAGCCCTGCGGAAATCCGTATGCTGAAG---TTGAACACCCAAACCTCGTGA 57
QY 195 CCTCATCGAGGTGTTTCAGGAGAAAAGGAAAATGCAATTTAGTATTTTGAATACCTGATCA 254
Db 58 CCTCATCGAGGTGTTTCAGGAGAAAAGGAAAAGATGCATCTAGTCTTTTGAGTACCTGATCA 117
QY 255 TACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGATGGAGTGATCAAAAG 314
Db 118 CACACTGTTAAACGAGCTGGAGAGAAACCCAAACCGGAGTTCTGATGGAGTGATTAAG 177
QY 315 CGTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGTCATATACATAAATGTTATTCACAG 374
Db 178 TGTGCTATGGCAAAACCTTCAAGCCCTTAACCTTCTGTCAAGGCACAAATTTGTTATTCATCG 237
QY 375 AGATATAAAACCTGAAATATTTCTAATAACTAAGCAAGGAAATTAATCAAGATTTGTGACTT 434
Db 238 GGATGTAACCTGAAACATCTCTAATAACCAAGCAAGGGATGATAAAGATTTGTGACTT 297
QY 435 CGGGTTTGCACAAATCTGATTCAGGAGATGCCTACACCGATTAATGTAGCTCAGATATG 494
Db 295 TGGATTTGCACGAATCTAATTCAGGAGACGCCTACACAGACTATGTTGCCACCGGTG 357
QY 495 GTACCGAGCTCTGAACCTTCTTGTGGAGATACTCAGTATGGTTCTTTCAGTCGATATG 554
Db 358 GTACCGAGCCCCGGAACCTTCTCGTGGAGACACGAAGTACGTTCTCTGTAGACGTG 417
QY 555 GGCTATTGGTGTGTTTTTTCAGAGCTCCTGACAGGCCAGCCACTGTGGCCTGGAAATC 614
Db 418 GGCCGTGGCTGTGTTTTTTCAGAG-CTCCTGACGGGTGAG-CACCTGTGGCCGGGAAACC 475

RESULT 7
AV986182
LOCUS AV986182 598 bp mRNA linear EST 14-MAR-2002
DEFINITION AV986182 Nori Satoh unpublished cDNA library, larva Ciona intestinalis cDNA clone cili4b18 5', mRNA sequence.
ACCESSION AV986182
VERSION AV986182.1 GI:19475050
KEYWORDS EST.
SOURCE Ciona intestinalis
ORGANISM Ciona intestinalis

REFERENCE Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona; Phlebobranchia; Cionidae; Ciona.
AUTHORS 1 (bases 1 to 598)
Satoh,N., Satou,Y., Kohara,Y. and Shin-i,T.
TITLE Expressed genes in Ciona intestinalis
JOURNAL Unpublished
COMMENT Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoh@ascidian.zool.kyoto-u.ac.jp.

FEATURES
source
1..598
/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"

```

/clone="c1lv4lb18"
/tissue_type="whole animal"
/dev_stage="larva"
/clone_lib="Nori Satch unpublished cDNA library, larva"
BASE COUNT      179 a   105 c   149 g   165 t
ORIGIN

Query Match      29.5%; Score 306.8; DB 9; Length 598;
Best Local Similarity 70.7%; Pred. No. 2.7e-11;
Matches 423; Conservative 0; Mismatches 172; Indels 3; Gaps 1;

QY 11 ATGAAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAAATGCAGAAACA 70
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1  ATGAGAAGATAGGTAAGATAGGGAAGGATCTTATGGCATTTGTTTAAATGTCGGAACC 60

QY 71 AAACCTCTGGACAAGTAGTAGCTGTAAAAAATTTGTGAAATCTGAAGATGATCCTGTTG 130
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 GTGACTCGGACAAATTGTGGCCATCAAAAATTTGTGAGTCAGAAGATGACCCGCTGA 120

QY 131 TTAAGAAAATAGCACTAAGAGAAATACGTATGTTGAGCAATTAACACATCCAAATCTTG 190
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 TCAAGAAAATCGCGCTGAGGGAATCAGAAATGTTAAAGCAATTGAAACACAAATACTTG 180

QY 191 TGAACCTCATCGAGGTGTTTCAGAGAAAAGGAAAATGCAATTTAGTTTTTGAATACTGTG 250
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 TGAACCTGATTGAAGTTTTCCGTCGAAAACGCAATTAACATCTTGTGTTCCGATATTG 240

QY 251 ATCATACACTTTTAAATGAGCTGGAAAGAAACCCCAATGGAGTTGCTGATGGAGTGATCA 310
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 ATCATACAGTTTTTAAATGAATTGGACAAACACATGAGAGGAGTACCAGAAACATCATGTTA 300

QY 311 AAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCATATACATACTGTTATTC 370
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 AACGTATTGTTGGCAAGTTCTTCAAGCTGTTTCACITTTGCCATCAGCATTAATTGTATAC 360

QY 371 ACAGAGATATAAAACCTGAAATATTCTAATAACTAAGCAAGGAATAATCAAGATTGTG 430
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 ATCGTGATGTAAACCAGAGAACATATTAAATCACTAAGCAAGGTGTTATCAAGTTGTG 420

QY 431 ACTTCGGTTTGCACAAATTCGA---TTCCAGGAGATGCCTACACCGATTATGTAGCTA 487
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 421 ACTTTGGATTGCAAGAATTTTAACTGGTCTCTGGCGAGATTACACCGACTATGTGGCCA 480

QY 488 CGAGATGGTACCGAGCTCCTGAACCTTCTTGTTGGAGATACTCAGTATGTTCTTCAGTGG 547
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 481 CTAGGTGTACAGAGACACCTGAGCTCCTCGTCGGGGACACTCAGTATGGCCCTCCAGTGG 540

QY 548 ATATATGGCTATTGGTTGTGTTTTCAGAGCTCTCGACAGGCCCGACCTGTGTGCC 605
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 541 ATGTGTGGGCAATAGGTTGCGTATTGCGGATTTGCGGATTTGTCGGGACAGGCCATTATGCC 598

RESULT 8
BU221831
LCCUS
DEFINITION      BU221831 824 bp mRNA linear EST 25-NOV-2002
                  603750354f1 CSEQCHN04 Gallus gallus cDNA clone CHEST661113 5', mRNA
                  sequence.
ACCESSION      BU221831
VERSION        BU221831.1 GI:25410266
KEYWORDS       EST.
SOURCE         Gallus gallus (chicken)
ORGANISM       Gallus gallus
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
                Phasianinae; Gallus.
REFERENCE      1 (bases 1 to 824)
AUTHORS        Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
                Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
TITLE          A Comprehensive Collection of Chicken cDNAs
JOURNAL        Curr. Biol. 12 (22), 1965-1969 (2002)
MEDLINE        22335534
PUBMED         12445392
COMMENT        Contact: Simon Hubbard
```

```

Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology (UMIST)
,
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers
1..824
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="CHEST661113"
/tissue_type="whole embryo"
/dev_stage="20-21"
/lab_host="DH10B"
/clone_lib="CSEQCHN04"
/note="Organ: whole embryo; Vector: pBluescript II KS(+);
Site 1: EcoRI; Site 2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunted, ligated to NotI adapters, digested with EcoRI
, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."
BASE COUNT      192 a   227 c   239 g   166 t
ORIGIN

Query Match      28.7%; Score 298.4; DB 13; Length 824;
Best Local Similarity 68.7%; Pred. No. 7.5e-11;
Matches 426; Conservative 0; Mismatches 191; Indels 3; Gaps 1;

QY 40 TCTTATGGGTTGTATTCAAATGCAGAAACAAAACCTCTCGACAAGTAGTAGTGTAA 99
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2  TCCTACGGCGTCTGCTTCAAGTGCCTCAACAGGAGACGGGGCAGATCGTGGCCATCAAG 61

QY 100 AAATTTGTGGAATCTGAAGATGATCCTGTTGTTAAGAAAATAGCACTAAGAGAAATACGT 159
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 62 AAGTTCCTGGAGTCCGAGGAGGACCCCGTGATCCGGAAGATCGCGTCCGGAGGTCCGC 121

QY 160 ATGTTGAAGCAATTAAAAACATCCAAATCTTGTGAACCTCATCGAGGTGTTCAAGGAAAA 219
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 122 ATGCTGAAGCAACTGAACACACCCCAACCTGTTGAACCTGCTGGAGGTGTTTCAGGAGAA 181

QY 220 AGGAAAATGCATTAGTTTTTGAATACTGTGATCATACACTTTTAAATGAGCTGGAAGA 279
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 182 AGGAAGCTGCACCTGGTCTTTGAGTACTGTGACCACACCCGTTCTCCACGAGCTGGACA 241

QY 280 AACCCAAATGGAGTTGCTGATGGAGTGATCAAAAGCGTATTATGGCAACACACTTCAAGCT 339
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 242 CACCCCGGGGGTCCCGGAGCAGCTGTCAGGAGCATTTACCTGGCAGACCCCTCCAACT 301

QY 340 CTTAATTTCTGTCATATACATAACTGATTATTCACAGAGATATAAAACCTGAAAATATTCTA 399
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 302 GTGAACCTTTTGTACAAAACACAACCTGCATCCATCGAGATGTAAAGCCAGAAAAACATCT 361

QY 400 ATAACCTAAGCAAGGAATAATCAAGATTTGTGACTTCGGGTTTGCACAAATTTCTGATT--- 456
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 362 ATAACGAAGCACTCCATCATCAAACTCTGTGACTTCGGATTTCGTCATGCTGACTGGC 421

QY 457 CCAGGAGATGCTTACACCGATTATGTAGCTACGAGATGGTACCGAGCTCCTGAACCTTCTT 516
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 422 CCAGGTGATTACTACACAGACTACGTGGCAACCCAGGTGTTGTTACCGTCCCCAGAGCTGCT 481

QY 517 GTGGGAGATACTCAGTATGGTTCTTCAGTCGATATATCGGCTATTGTTGTTTTTGCA 576
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```


SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
AUTHORS ; Rhabditidae; Peloderinae; Caenorhabditis.
1 (bases 1 to 565)
Reboul,J., Vaglio,P., Rual,J.F., Lamesch,P., Martinez,M., Armstrong
C.M., Li,S., Jacotot,L., Bertin,N., Janky,R., Moore,T., Hudson
J.R., Hartley,J.L., Brasch,M.A., Vandenhaute,J., Boulton,S.,
Endress,G.A., Jenna,S., Chevet,E., Papasotiropoulos,V., Collas,P.P.,
Ptacek,J., Snyder,M., Huang,R., Chance,M.R., Lee,H.,
Doucette-Stamm,L., Hill,D.E. and Vidal,M.
C. elegans ORFeome version 1.1: experimental verification of the
genome annotation and resource for proteome-scale protein
expression
JOURNAL Nat. Genet., (2003) In press
COMMENT Contact: Vidal M
Marc Vidal Laboratory
Dana Farber Cancer Institute
1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
Tel: 617 632 5180
Fax: 617 632 5739
Email: Marc.Vidal@dfci.harvard.edu
Sequence tag of Gateway entry clones. The primers used were
designed on the predicted protein encoding ORF. C. elegans ORFeome
cloning project : Contact david_hill@dfci.harvard.edu or
marc_vidal@dfci.harvard.edu
POLYA=No.

FEATURES
source
1..565
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/sex="Hermaphrodite and male"
/tissue_type="whole animal"
/dev_stage="mixed stage"
/clone_lib="AD-wrmCDNA"
/note="The AD-wrmCDNA library was generated with poly(A)
RNA isolated from both hermaphrodite and male N2 worms of
all larval stages, embryos, adults and dauers and the
subsequent generation of cDNAs by poly(A) priming. The
cDNAs were cloned into pPC86"
BASE COUNT 184 a 88 c 135 g 158 t

Query Match 25.9%; Score 269.8; DB 14; Length 565;
Best Local Similarity 67.6%; Pred. No. 5.7e-09;
Matches 379; Conservative 0; Mismatches 182; Indels 0; Gaps 0;

QY 31 GGAGAAGGTCTTATGGGTTGTATCAAAATGCAGAAACAAACCTCTGGACAAGTAGTA 90
DB 5 GGTGAAGTTCATATGGAGTTGTATATAAATGTAATAATAGGATACCTGGACAATTTGTG 64
QY 91 GCTGTTAAAAAATTTGTGGAATCTGAAGATGATCCTGTGTTTAAAGAAATAGCACTAAGA 150
DB 65 GCAATCAAAAAGTTTGTGGAACCTGAAGATGATCCGCATATTAAGAAAATTCGCTACGG 124
QY 151 GAAATACGTATGTTGAAGCAATTAACATCCAAATCTGTGAACCTCATCGAGGTGTTT 210
DB 125 GAAATCAGAAATGCTGAAGCAACTGAACATCAAAATTTGTTGGATTGATTGAAGTGTTC 184
QY 211 AGGAGAAAAGAAAATGCAATTTAGTTTGAATCTGTGATCATACACTTTTAAATGAG 270
DB 185 AAACGGAACCGAAACTTCATCTCGTCTTGAACCTTTGTGATCGAAGTGTACTTCACGAG 244
QY 271 CTGGAAGAAACCCAAATGGAGTTGCTGATGAGTGAAGGCGTATTATGCAACAA 330
DB 245 TTGGAGAAGAAATCCGCATGGAGTTAACCATGAACCTCATAAAGAAAATAATTATCAATTG 304
QY 331 CTTCAGCTCTTAATTTCTGTCATATACATAACTGTATTTCACAGAGATATAAACCTGAA 390
DB 305 CTAGAAGCTCTTAATTTCTGTCACAGTCACAAATGTATTTCATCGAGATGTGAACCGGAA 364

QY 391 AATATTCTAATAACTAAGCAAGGAATAATCAAGATTTGTGACTTCGGGTTTGCACAAAATT 450
DB 365 AACATTTTCTTGACACGGAATGATCAAGTGAAACTTTGGAGATTTCCGATTTGCTCGAATA 424
QY 451 CTGATTCAGAGAGATGCCTACACCGATTATGTAGCTACGAGATGGTACCGAGCTCCTGAA 510
DB 425 ATAAACACGACGGAATGTATACTGACTATGTGGCAACTCGTTGGTACCGTACTCCAGAA 484
QY 511 CTTCTTGTGGGAGATACACTCAGTATGGTTCTTCAGTCGATATATGGCTATTGGTTGTGT 570
DB 485 CTTCTGGTGGTGTATGTTCAATATGGGCCACCTGTAGACATTTGGGCTGTAGGATGTGTA 544
QY 571 TTTGCAGAGCTCCTGACAGGC 591
DB 545 TATGGGAGCTGTTGACTGGC 565

RESULT 13
LOCUS AW106692
DEFINITION um32d03.y1 Sugano mouse kidney mkia Mus musculus cDNA clone
IMAGE:2236229 5' similar to gb:X66358 SERINE/THREONINE-PROTEIN
KINASE KKIALRE (HUMAN);, mRNA sequence.
ACCESSION AW106692
VERSION AW106692.1 GI:6077492
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 731)
REFERENCE Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter
E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
Waterston,R. and Wilson,R.
The WashU-NCI Mouse EST Project 1999
Unpublished
Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNC; ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
XGI:1006441
Seq primer: custom primer used
High quality sequence stop: 433.

FEATURES
source
1..731
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:2236229"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Sugano mouse kidney mkia"
/note="Organ: kidney; Vector: pME18S-FL3; Site 1: DraIII
(CACTGTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGCGCTTTTCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor [TGTGGCCCTACTGG]; digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTCTGCTCTAAAGCTGCG and 3' end
primer CGACCTGCAGCTCGAGCACA."

Db 647 TCCCCCTGTAGATGCTGGGC-AATGGCTGTGTGTTTGCTGAGTTACTGTCCGGAGTGCC 705

Qy 597 ACTGTGGCTTGGGA 609

Db 706 TCTATGGCAAGGA 718

RESULT 15

BI511582

LOCUS

DEFINITION

BB160006A20C12.5 Bee Brain Normalized Library, BB16 Apis mellifera

CDNA clone BB160006A20C12 5', mRNA sequence.

BI511582

VERSION

BI511582.1 GI:15361956

KEYWORDS

EST.

SOURCE

Apis mellifera (honeybee)

ORGANISM

Apis mellifera

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea; Apidae; Apis.

REFERENCE

1 (bases 1 to 508)

AUTHORS

Whitfield,C.W., Band,M.R., Bonaldo,M.F., Kumar,C.G., Liu,L., Pardin,J., Robertson,H.M., Soares,B. and Robinson,G.E.

TITLE

Annotated expressed sequence tags and cDNA microarrays for studies of brain and behavior in the honey bee

JOURNAL

Genome Res. 12 (4), 555-566 (2002)

MEDLINE

21929762

PUBMED

11932240

COMMENT

Contact: Gene E. Robinson

Department of Entomology

University of Illinois

505 S. Goodwin Ave., Urbana, IL 61801, USA

Tel: 217 265 0309

Fax: 217 244 3499

Email: generobi@life.uiuc.edu

This research was funded by the University of Illinois Critical Research Initiatives Fund and a Burroughs-Wellcome Trust Innovation Award in Functional Genomics to G.E. Robinson and an NSF Postdoctoral Fellowship in Bioinformatics to C.W. Whitfield.

PCR Primers

FORWARD: TAATACGACTCACTATAGGG

BACKWARD: ATTAACCTCACTAAAG

Plate: BB160006A20 row: C column: 12

Seq primer: AGCGGATAACAATTTCACACAGGA

High quality sequence stop: 508.

Location/Qualifiers

1..508

/organism="Apis mellifera"

/mol_type="mRNA"

/strain="mixed strains of European bees, predominantly A.m. ligustica"

/db_xref="taxon:7460"

/clone="BB160006A20C12"

/sex="female"

/tissue_type="brain"

/dev_stage="adult worker honey bee"

/lab_host="DH10B"

/clone_lib="Bee Brain Normalized Library, BB16"

/note="Organ: Brain; Vector: pTT3-pac; Site 1: EcoRI; Site 2: NotI; The BB16 library was contributed by the Soares laboratory and it was constructed and normalized as described by Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome Research 6(9): 791-806. RNA was prepared from dissected brains of adult worker bees of various ages and various behavioral groups."

BASE COUNT

163 a 93 c 118 g 134 t

ORIGIN

Query Match 23.9%; Score 249; DB 12; Length 508;

Best Local Similarity 68.3%; Pred. No. 1.2e-37;

Matches 345; Conservative 0; Mismatches 160; Indels 0; Gaps 0;

QY 110 AATCTGAGATGATCCTGTTGTTAAGAAATAGCACTAGAGAAATACGTTATGTAAGC 169

Db 4 AAACAGAGATGATCCTCTTATACGCAAAATTCGCTACGTGAATAAGACTTCTTAAGA 63

Qy 170 AATTAAACATCCAAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAGGAAATGC 229

Db 64 ATCTTAACATCCAAATCTGTGCAATCTTCTGGAGGCTTCAGACGGAACCGAAATTAC 123

Qy 230 ATTTAGTTTTGAATACTGTGTATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATG 289

Db 124 ACCTAGTCTTTGAATATTGCGAGTATACGTTATTAAATGAGATGGAGAGATATCCGAGTG 183

Qy 290 GAGTTGCTGATGGAGTGATCAAAAGCGTATTATGGCAACACACTTCAAGCTCTTAATTTCT 349

Db 184 GGTGTCCGGAATCACTACAAACAACTCACATGGGCAATACTGCAGGGCATCGCATATT 243

Qy 350 GTCATATACATAACTGTATTTCAGAGATATAAAACCTGAAAATATTCTAATAACTAAGC 409

Db 244 GCCATCGATTGGGTTGTTCATAGAGATGTAAAGCCGGAATATTTTAAATCACAGGCG 303

Qy 410 AAGGAATATCAAGATTTGTGACTTCGGGTTTGCACAAATTTCTGATTCACAGGAGATGCCT 469

Db 304 ATGGCGTGTAAATTTGTGTGACTTCGGTTTCGCTCGAATGCTCAGTCTCTGGTGAATTT 363

Qy 470 ACACCGATTATGTAGCTACGAGATGGTACCGAGCTCCTGAACCTTCTTGTGGGAGATACTC 529

Db 364 ACACAGAATATGTTGCAACTAGATGGTACAGAGCACCTGAACCTATTGGTTGGAGATACTC 423

Qy 530 AGTATGGTTCTTCAGTCGATATATGGGCTATTGGTTGTTTTCAGAGCTCCTCGACAG 589

Db 424 AATACGGTACACCAAGTCGATGTATGGGCAATCGGTTGTTTTCGAGAACTAATACGTTG 483

Qy 590 GCCAGCCACTGTGGCCTGGAATAATC 614

Db 484 GAGAAGCACTATGGCCAGGAAATC 508

Search completed: November 15, 2003, 20:27:44

Job time : 2627.85 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2003, 11:38:36 ; Search time 327.082 Seconds
(without alignments)
8591.481 Million cell updates/sec

Title: US-09-671-050-9
Perfect score: 1041
Sequence: 1 atggaaaagtatgaaaatt.....aggtacttcggtcaaaagt 1041

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03.*
1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1041	100.0	1041	22	AAD03816 Human kinase cDNA
2	1025	98.5	1068	22	AAD03813 Human kinase cDNA
3	839	80.6	945	22	AAD03817 Human kinase cDNA
4	823	79.1	972	22	AAD03814 Human kinase cDNA
5	821	78.9	1790	24	AAD30557 Human kinase polyp
6	762.8	73.3	1678	25	ABX08936 cDNA encoding huma
7	757	72.7	1083	22	AAS06725 Polynucleotide seq
8	709	68.1	1819	22	AAD03818 Human kinase cDNA

9	644	61.9	882	24	AAI64248 Human kinase 14257
10	543.2	52.2	1281	25	ABX34679 Human mddt cDNA SE
11	537	51.6	911	22	AAF44669 Novel protein kina
12	454.4	43.7	561	22	AAD03812 Human kinase cDNA
13	454.4	43.7	594	22	AAD03815 Human kinase cDNA
14	453.2	43.5	2615	22	AAF44670 Novel protein kina
15	395	37.9	1086	25	ABZ77165 Human protein kina
16	343.8	33.0	1612	24	ABQ93433 Human cDNA SEQ ID
17	343.8	33.0	2944	23	ABV24584 Human prostate exp
18	251.8	24.2	1698	21	AAA29746 Rabbit KKIAMRE kin
19	251.8	24.2	3080	21	AAA29745 Rabbit KKIAMRE kin
20	242.2	23.3	1179	23	ABL12603 Drosophila melanog
21	202.6	19.5	1513	21	AAS1208 Human lost in leuk
22	202.6	19.5	1773	22	AAS06724 Polynucleotide seg
23	191	18.3	1667	21	AAZ51207 Rat lost in leukae
24	191	18.3	1738	21	AAZ51206 Rat lost in leukae
25	186.6	17.9	3297	23	ABL12602 Drosophila melanog
26	177.4	17.0	1158	23	AAS70329 DNA encoding novel
27	175.2	16.8	1866	22	AAF44672 Novel protein kina
28	163.8	15.7	903	20	AAZ07476 Pneumocystis carin
29	162	15.6	1476	21	AAZ37835 Cyclin-dependent k
30	162	15.6	1635	19	AAV71073 CDK2-green fluores
31	162	15.6	1635	19	AAV71074 Green fluorescent
32	161.6	15.5	1024	25	ABZ83372 Toxicologically re
33	161.6	15.5	1297	24	ABQ60787 Human 10ckshs1 cDN
34	160.4	15.4	2213	24	ABK84073 Human cDNA differe
35	160	15.4	1322	21	AAF15889 Human prostate can
36	154.2	14.8	1050	24	ABT11085 Human breast cance
37	154.2	14.8	1050	24	ABL64774 Lung cancer relate
38	154.2	14.8	1050	24	ABL65408 Lung cancer relate
39	154.2	14.8	1050	24	ABL65855 Lung cancer relate
40	154.2	14.8	1050	24	ABL69118 Kidney cancer rela
41	154.2	14.8	1050	25	ACC50096 Breast cancer asso
42	154.2	14.8	1050	25	AAD52783 Human CDC2 DNA. H
43	154.2	14.8	1235	25	ABX76248 Lung cancer-associ
44	154.2	14.8	1825	24	AAS94983 Human DNA sequence
45	154.2	14.8	1897	25	ACA03966 cDNA downregulated

ALIGNMENTS

RESULT 1
AAD03816
ID AAD03816 standard; cDNA; 1041 BP.
XX
AC AAD03816;
XX DT 19-JUN-2001 (first entry)
DE Human kinase cDNA #5.
XX KW Human; kinase; gene therapy; bioreactor; mental disorder;
KW biological disorder; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT CDS 1..1041
FT /*tag= a
FT /product= "Human kinase #5"
FT /note= "The coding region does not include stop codon."
XX WO200123579-A1.
XX PD 05-APR-2001.
XX PF 27-SEP-2000; 2000WO-US26621.
XX PR 28-SEP-1999; 99US-0156511.
XX PA (LEXI-) LEXICON GENETICS INC.

XX Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
PI WPI; 2001-266166/27.
XX P-PSDB; AAE00494.
DR
XX New isolated human kinase polynucleotide useful for generating
PT antibodies, as reagents in diagnostic assays and for screening for
PT compounds useful for treating mental, biological or medical diseases
XX
XX Claim 1; Page 31-32; 38pp; English.
PS
XX The present sequence is a cDNA encoding novel human protein
CC (NHP) known as human kinase. The human kinases share structural
CC similarity with animal kinases, more particularly serine or
CC threonine protein kinases. Human kinase cDNA is useful for the
CC detection of mutant human kinase for the diagnosis of disease,
CC and also as a therapeutic. It is useful for screening drugs
CC effective in the treatment of symptomatic or phenotypic
CC manifestations perturbing the normal function of NHP in the
CC body. The NHP nucleotide sequences are useful for generation of
CC antibodies, as reagents in diagnostic assays, for the
CC identification of other cellular gene products related to human
CC kinases, and as reagents in assays for screening compounds that
CC are useful for treating mental, biological or medical disorders.
CC NHP oligonucleotides are used as probes. The labelled NHP probes
CC are useful for screening human genomic library for identifying
CC polymorphisms and as primers in amplification assays to detect
CC mutations within the exons, introns and splice sites that can
CC be used in diagnostics and pharmacogenomics. Nucleotide construct
CC encoding NHP products are used to genetically engineer cells
CC in vivo that functions as bioreactors in the body delivering a
CC continuous supply of NHP to the body. Nucleotide constructs
CC encoding functional NHPs are used in gene therapy for the
CC modulation of NHP expression.
XX
XX Sequence 1041 BP; 344 A; 190 C; 228 G; 279 T; 0 other;
SQ
Query Match 100.0%; Score 1041; DB 22; Length 1041;
Best Local Similarity 100.0%; Pred. No. 1.5e-268;
Matches 1041; Conservative C; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAAAGATGATAAAATAGCTAAGACTGGAGAGGGTCTTATGGGGTGTATTCAA 60
DB 1 ATGGAAAGATGATAAAATAGCTAAGACTGGAGAGGGTCTTATGGGGTGTATTCAA 60
QY 61 TGCAGAAACAAACCTCTGGCAAGTAGTAGCTGTTAAATAATTTGTGGAATCTGAAGAT 120
DB 61 TGCAGAAACAAACCTCTGGCAAGTAGTAGCTGTTAAATAATTTGTGGAATCTGAAGAT 120
QY 121 GATCCTGTGTTAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAT 180
DB 121 GATCCTGTGTTAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAT 180
QY 181 CCAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAGGAAATGCAATTTAGTTT 240
DB 181 CCAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAGGAAATGCAATTTAGTTT 240
QY 241 GAATCTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAATTTGCTGAT 300
DB 241 GAATCTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAATTTGCTGAT 300
QY 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCATATACAT 360
DB 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCATATACAT 360
QY 361 AACTGTATTACAGAGATATAAAACCTGAAATATTCTAATAACTAAGCAGGAATAATC 420
DB 361 AACTGTATTACAGAGATATAAAACCTGAAATATTCTAATAACTAAGCAGGAATAATC 420
QY 421 AAGATTGTGACTTCGGGTTTGGCAAAATTTCTGATTCAGGAGATGCCTACACCGATTAT 480
DB 421 AAGATTGTGACTTCGGGTTTGGCAAAATTTCTGATTCAGGAGATGCCTACACCGATTAT 480

QY 481 GTAGCTACGAGATGGTACCGAGCTCCTGAACCTTCTTGTGGGAGATACTCAGTATGGTTCT 540
DB 481 GTAGCTACGAGATGGTACCGAGCTCCTGAACCTTCTTGTGGGAGATACTCAGTATGGTTCT 540
QY 541 TCAGTCGATATATGGGCTATTGGTTGTGTTTTCAGAGAGCTCCTGACAGGCCAGGCACTG 600
DB 541 TCAGTCGATATATGGGCTATTGGTTGTGTTTTCAGAGAGCTCCTGACAGGCCAGGCACTG 600
QY 601 TGGCCTGGAAATCAGATGTGGACCAACTTTTATCTGATAATCAGAACACTAGTAGAGACG 660
DB 601 TGGCCTGGAAATCAGATGTGGACCAACTTTTATCTGATAATCAGAACACTAGTAGAGACG 660
QY 661 GGGTTTCGCCATGTTGACCCAGGCTGGTCTCGAACTCTTGACGTCAAGTGATCCACCTGCC 720
DB 661 GGGTTTCGCCATGTTGACCCAGGCTGGTCTCGAACTCTTGACGTCAAGTGATCCACCTGCC 720
QY 721 GTAGCCTCTCAAAGTCTGGAATTACAGGAAATTAATCCCAAGACATCAATCAATCTTT 780
DB 721 GTAGCCTCTCAAAGTCTGGAATTACAGGAAATTAATCCCAAGACATCAATCAATCTTT 780
QY 781 AAAAGTACCGGTTTTCCTATGGCATCAGTATACCTGAGCCAGAGACATGGAAACTCTT 840
DB 781 AAAAGTACCGGTTTTCCTATGGCATCAGTATACCTGAGCCAGAGACATGGAAACTCTT 840
QY 841 GAGGAAAAGTTCTCAGATGTTTCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
DB 841 GAGGAAAAGTTCTCAGATGTTTCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
QY 901 ATGAATCCAGATGACAGATTTAACTGTTTCCCACTCTGAGAGAGCTCCTACTTTGATTCT 960
DB 901 ATGAATCCAGATGACAGATTTAACTGTTTCCCACTCTGAGAGAGCTCCTACTTTGATTCT 960
QY 961 TTTCAAGAGGCCCAAATTTAAAGAAAGACGTAATGAAGGAAGAAACAGAGACGCCAA 1020
DB 961 TTTCAAGAGGCCCAAATTTAAAGAAAGACGTAATGAAGGAAGAAACAGAGACGCCAA 1020
QY 1021 CAGGTACTTCGCTCAAAAGT 1041
DB 1021 CAGGTACTTCGCTCAAAAGT 1041

RESULT 2
AAD03813
ID AAD03813 standard; cDNA; 1068 BP.
XX
AC AAD03813;
XX
DT :9-JUN-2001 (first entry)
XX
DE Human kinase cDNA #2.
XX
KW Human; kinase; gene therapy; bioreactor; mental disorder;
KW biological disorder; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1068
FT /*tag= a
FT /product= "Human kinase #2"
FT /note= "The coding region does not include stop codon"
FT /partial
XX
PN WO200123579-A1.
XX
PD 05-APR-2001.
XX
PF 27-SEP-2000; 2000WO-US26621.
XX
PR 28-SEP-1999; 99US-0156511.
XX
PA (LEXI-) LEXICON GENETICS INC.

XX Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
PI WPI; 2001-266166/27.
XX P-PSDB; AAE00491.
PT New isolated human kinase polynucleotide useful for generating
PT antibodies, as reagents in diagnostic assays and for screening for
PT compounds useful for treating mental, biological or medical diseases -
XX Disclosure; Page 28; 38pp; English.
XX The present sequence is a cDNA encoding novel human protein
CC (NHP) known as human kinase. The human kinases share structural
CC similarity with animal kinases, more particularly serine or
CC threonine protein kinases. Human kinase cDNA is useful for the
CC detection of mutant human kinase for the diagnosis of disease,
CC and also as a therapeutic. It is useful for screening drugs
CC manifestations perturbing the normal function of NHP in the
CC body. The NHP nucleotide sequences are useful for generation of
CC antibodies, as reagents in diagnostic assays, for the
CC identification of other cellular gene products related to human
CC kinases, and as reagents in assays for screening compounds that
CC are useful for treating mental, biological or medical disorders.
CC NHP oligonucleotides are used as probes. The labelled NHP probes
CC are useful for screening human genomic library for identifying
CC polymorphisms and as primers in amplification assays to detect
CC mutations within the exons, introns and splice sites that can
CC be used in diagnostics and pharmacogenomics. Nucleotide construct
CC encoding NHP products are used to genetically engineer cells
CC in vivo that functions as bioreactors in the body delivering a
CC continuous supply of NHP to the body. Nucleotide constructs
CC encoding functional NHPs are used in gene therapy for the
CC modulation of NHP expression.
XX Sequence 1068 BP; 344 A; 197 C; 240 G; 287 T; 0 other;
SQ Query Match 98.5%; Score 1025; DB 22; Length 1068;
Best Local Similarity 100.0%; Pred. No. 2.9e-264;
Matches 1025; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGAAAAGTATGAAAATAGCTAGACTGGAGAGGGTCTTATGGGTTGTATTCAA 60
DB 1 ATGGAAAAGTATGAAAATAGCTAGACTGGAGAGGGTCTTATGGGTTGTATTCAA 60
QY 61 TGCAGAAACAAACCTCTGGACAAGTAGCTGTTAAAAAATTTGTGGAATCTGAAGAT 120
DB 61 TGCAGAAACAAACCTCTGGACAAGTAGCTGTTAAAAAATTTGTGGAATCTGAAGAT 120
QY 121 GATCCTGTTGTAAGAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAT 180
DB 121 GATCCTGTTGTAAGAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAT 180
QY 181 CCAAACTCTTGAAACCTCATCGAGGTGTTCAAGGAGAAAAGGAAAATGCATTTAGTTT 240
DB 181 CCAAACTCTTGAAACCTCATCGAGGTGTTCAAGGAGAAAAGGAAAATGCATTTAGTTT 240
QY 241 GAATACCTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT 300
DB 241 GAATACCTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT 300
QY 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTATATACAT 360
DB 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTATATACAT 360
QY 361 AACTGTATTACAGAGATATATAAACCTGAAAATATTCTAATAACTAAGCAAGGAATAATC 420
DB 361 AACTGTATTACAGAGATATATAAACCTGAAAATATTCTAATAACTAAGCAAGGAATAATC 420
QY 421 AAGATTGTGACCTCGGGTTTGCACAAATTTCTGATTCAGGAGATGCTTACACCGATTAT 480
DB 421 AAGATTGTGACCTCGGGTTTGCACAAATTTCTGATTCAGGAGATGCTTACACCGATTAT 480

QY 481 GTAGCTACGAGATGGTACCGAGCTCCTGAACCTTCTTGTGGAGATATCAGTATGGTTCT 540
DB 481 GTAGCTACGAGATGGTACCGAGCTCCTGAACCTTCTTGTGGAGATATCAGTATGGTTCT 540
QY 541 TCAGTCGATATATGGCTATTTGGTTGTGTTTTCAGAGCTCCTGACAGGCCAGCCACTG 600
DB 541 TCAGTCGATATATGGCTATTTGGTTGTGTTTTCAGAGCTCCTGACAGGCCAGCCACTG 600
QY 601 TGGCCTGGAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACTAGTAGAGAG 660
DB 601 TGGCCTGGAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACTAGTAGAGAG 660
QY 661 GGGTTTCGCCATGTTGACCAAGGCTGGTCTCGAACTCTTGACGTCAAGTGATCCACCTGCC 720
DB 661 GGGTTTCGCCATGTTGACCAAGGCTGGTCTCGAACTCTTGACGTCAAGTGATCCACCTGCC 720
QY 721 GTAGCCTCTCAAAGTGTCTGGAATTACAGGAAAATTAATCCCAAGACATCAATCAATCTTT 780
DB 721 GTAGCCTCTCAAAGTGTCTGGAATTACAGGAAAATTAATCCCAAGACATCAATCAATCTTT 780
QY 781 AAAAGTAACGGGTTTTTCCATGGCATCAGTATACCTGAGCCAGAGACATGGAAACTCTT 840
DB 781 AAAAGTAACGGGTTTTTCCATGGCATCAGTATACCTGAGCCAGAGACATGGAAACTCTT 840
QY 841 GAGGAAAAGTTCTCAGATGTTTCATCCTGTGGCTCTGAACTTCATGAAGGGGTGCTGAAG 900
DB 841 GAGGAAAAGTTCTCAGATGTTTCATCCTGTGGCTCTGAACTTCATGAAGGGGTGCTGAAG 900
QY 901 ATGAATCCAGATGACAGATTAACTGTTTCCCAACTCCTGGAGAGCTCTTACTTTGATTCT 960
DB 901 ATGAATCCAGATGACAGATTAACTGTTTCCCAACTCCTGGAGAGCTCTTACTTTGATTCT 960
QY 961 TTTCAGAGGGCCCAATTAAAGAAAAGCAGTAAATGAAGGAAGAAAGAGAGAGCCCAA 1020
DB 961 TTTCAGAGGGCCCAATTAAAGAAAAGCAGTAAATGAAGGAAGAAAGAGAGAGCCCAA 1020
QY 1021 CAGGT 1025
DB 1021 CAGGT 1025
RESULT 3
AAD03817
ID AAD03817 standard; cDNA; 945 BP.
XX
AC AAD03817;
XX
DT 19-JUN-2001 (first entry)
XX
DE Human kinase cDNA #6.
XX
KW Human; kinase; gene therapy; bioreactor; mental disorder;
biological disorder; ss.
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..945
FT /tag= a
FT /product= "Human kinase #6"
FT /note= "The coding region does not include stop codon"
FT /partial
XX
PN W0200123579-A1.
XX
PD 05-APR-2001.
XX
PF 27-SEP-2000; 2000WO-US26621.
XX
PR 28-SEP-1999; 99US-0156511.
XX (LEXI-) LEXICON GENETICS INC.
PA

XX Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
P: WPI; 2001-266166/27.
XX P-PSDB; AAE00495.
XX
P: New isolated human kinase polynucleotide useful for generating
P: antibodies, as reagents in diagnostic assays and for screening for
P: compounds useful for treating mental, biological or medical diseases
XX
PS Claim 3; Page 33; 38pp; English.
XX
CC The present sequence is a cDNA encoding novel human protein
CC (NHP) known as human kinase. The human kinases share structural
CC similarity with animal kinases, more particularly serine or
CC threonine protein kinases. Human kinase cDNA is useful for the
CC detection of mutant human kinase for the diagnosis of disease,
CC and also as a therapeutic. It is useful for screening drugs
CC effective in the treatment of symptomatic or phenotypic
CC manifestations perturbing the normal function of NHP in the
CC body. The NHP nucleotide sequences are useful for generation of
CC antibodies, as reagents in diagnostic assays, for the
CC identification of other cellular gene products related to human
CC kinases, and as reagents in assays for screening compounds that
CC are useful for treating mental, biological or medical disorders.
CC NHP oligonucleotides are used as probes. The labelled NHP probes
CC are useful for screening human genomic library for identifying
CC polymorphisms and as primers in amplification assays to detect
CC mutations within the exons, introns and splice sites that can
CC be used in diagnostics and pharmacogenomics. Nucleotide construct
CC encoding NHP products are used to genetically engineer cells
CC in vivo that functions as bioreactors in the body delivering a
CC continuous supply of NHP to the body. Nucleotide constructs
CC encoding functional NHPs are used in gene therapy for the
CC modulation of NHP expression.
XX
SQ Sequence 945 BP; 323 A; 165 C; 202 G; 255 T; 0 other;

Query Match 80.6%; Score 839; DB 22; Length 945;
Best Local Similarity 90.8%; Pred. No. 1.7e-214;
Matches 945; Conservative 0; Mismatches 0; Indels 36; Gaps 1;

QY 1 ATGGAAAGTATGAAATATTAGCTAAGACTGGAGAGGGTCTTATGGGTGTATTCAA 60
DB 1 ATGGAAAGTATGAAATATTAGCTAAGACTGGAGAGGGTCTTATGGGTGTATTCAA 60
QY 61 TGCAGAAACAAACCTCTGGACAAGTAGTAGCTGTAAATAAATTTGTGGAACTGAAGAT 120
DB 61 TGCAGAAACAAACCTCTGGACAAGTAGTAGCTGTAAATAAATTTGTGGAACTGAAGAT 120
QY 121 GATCCTGTTGTTAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAT 180
DB 121 GATCCTGTTGTTAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAT 180
QY 181 CCAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAGGAAATGCAATTTATTTT 240
DB 181 CCAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAGGAAATGCAATTTATTTT 240
QY 241 GAATACCTGTGATCATACCTTTTANATGAGCTGGAAAGAAACCCAAATGAGTGTCTGAT 300
DB 241 GAATACCTGTGATCATACCTTTTAAATGAGCTGGAAAGAAACCCAAATGAGTGTCTGAT 300
QY 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGCATATACAT 360
DB 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGCATATACAT 360
QY 361 AACTGTATTCACAGAGATATAAACCTGAAATATTCTAATACTAAGCAAGGAATAATC 420
DB 361 AACTGTATTCACAGAGATATAAACCTGAAATATTCTAATACTAAGCAAGGAATAATC 420
QY 421 AAGATTTGTGACTTCGGGTTTGCACAAATTTCTGATTCAGGAGATGCCTACACCGATTAT 480
DB 421 AAGATTTGTGACTTCGGGTTTGCACAAATTTCTGATTCAGGAGATGCCTACACCGATTAT 480

QY 481 GTAGCTACGAGATGGTACCGAGCTCCTGAACTTCTTGTGGAGATACCTCAGTATGGTTCT 540
DB 481 GTAGCTACGAGATGGTACCGAGCTCCTGAACTTCTTGTGGAGATACCTCAGTATGGTTCT 540
QY 541 TCAGTCGATATATGGGCTATTGGTGTGTTTTTGCAGAGCTCCTGACAGGCCAGCCACTG 600
DB 541 TCAGTCGATATATGGGCTATTGGTGTGTTTTTGCAGAGCTCCTGACAGGCCAGCCACTG 600
QY 601 TGGCCTGGAAATCAGATGTGGACCAACTTTATCTGATAATCAGAAACACTAGTAGAGACG 660
DB 601 TGGCCTGGAAATCAGATGTGGACCAACTTTATCTGATAATCAGAAACACT----- 650
QY 661 GGGTTTCGCCCATGTTGACCAGGCTGGTCTCGAACTCTTGACGTCAAGTGATCCACCTGCC 720
DB 651 ----- 650
QY 721 GTAGCCTCTCAAAGTGCTGGAATTACAGGAAATTAATCCCAAGACATCAATCAATCTTT 780
DB 651 -----AGGAAATTAATCCCAAGACATCAATCAATCTTT 684
QY 781 AAAAGTAACGGGTTTTTCCATGGCATCAGTATACCTGAGCCAGAGACATGGAAACTCTT 840
DB 685 AAAAGTAACGGGTTTTTCCATGGCATCAGTATACCTGAGCCAGAGACATGGAAACTCTT 744
QY 841 GAGGAAAGTCTCAGATGTTTCATCCTGTGGCTCTGAACTTCATGAAGGGGTGCTGAAG 900
DB 745 GAGGAAAGTCTCAGATGTTTCATCCTGTGGCTCTGAACTTCATGAAGGGGTGCTGAAG 804
QY 901 ATGAATCCAGATGACAGATTTAACTTCCCAACTCCTGGAGAGCTCCTACTTTGATTCT 960
DB 905 ATGAATCCAGATGACAGATTTAACTTCCCAACTCCTGGAGAGCTCCTACTTTGATTCT 864
QY 961 TTTCAAGAGGGCCCAATTAAGAAAGAAAGCAACGTAATGAAGGAGAAACAGAGACGCCAA 1020
DB 865 TTTCAAGAGGGCCCAATTAAGAAAGAAAGCAACGTAATGAAGGAGAAACAGAGACGCCAA 924
QY 1021 CAGGTACTTCCGCTCAAAAGT 1041
DB 925 CAGGTACTTCCGCTCAAAAGT 945
RESULT 4
AADC3814
ID AADC3814 standard; cDNA; 972 BP.
XX
AC AADC3814;
XX
DT :9-JUN-2001 (first entry)
XX
DE Human kinase cDNA #3.
XX
KW Human; kinase; gene therapy; bioreactor; mental disorder;
KW biological disorder; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 1..972
FT /*tag= a
FT /product= "Human kinase #3"
FT /note= "The coding region does not include stop codon"
XX
PN WO200123579-A1.
XX
PD 05-APR-2001.
XX
PF 27-SEP-2000; 2000WO-US26621.
XX
PR 28-SEP-1999; 99US-0156511.
XX
PA (LEXI-) LEXICON GENETICS INC.

XX Doncho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
PI WPI; 2001-266166/27.
XX P-PSDB; AAE00492.
DR
DR
XX
PT New isolated human kinase polynucleotide useful for generating
PT antibodies, as reagents in diagnostic assays and for screening for
PT compounds useful for treating mental, biological or medical diseases -
XX
PS Disclosure; Page 29-30; 38pp; English.
XX
CC The present sequence is a cDNA encoding novel human protein
CC (NHP) known as human kinase. The human kinases share structural
CC similarity with animal kinases, more particularly serine or
CC threonine protein kinases. Human kinase cDNA is useful for the
CC detection of mutant human kinase for the diagnosis of disease,
CC and also as a therapeutic. It is useful for screening drugs
CC effective in the treatment of symptomatic or phenotypic
CC manifestations perturbing the normal function of NHP in the
CC body. The NHP nucleotide sequences are useful for generation of
CC antibodies, as reagents in diagnostic assays, for the
CC identification of other cellular gene products related to human
CC kinases, and as reagents in assays for screening compounds that
CC are useful for treating mental, biological or medical disorders.
CC NHP oligonucleotides are used as probes. The labelled NHP probes
CC are useful for screening human genomic library for identifying
CC polymorphisms and as primers in amplification assays to detect
CC mutations within the exons, introns and splice sites that can
CC be used in diagnostics and pharmacogenomics. Nucleotide construct
CC encoding NHP products are used to genetically engineer cells
CC in vivo that functions as bioreactors in the body delivering a
CC continuous supply of NHP to the body. Nucleotide constructs
CC encoding functional NHPs are used in gene therapy for the
CC modulation of NHP expression.
XX
SQ Sequence 972 BP; 323 A; 172 C; 214 G; 263 T; 0 other;

Query Match: 79.1%; Score 823; DB 22; Length 972;
Best Local Similarity 90.6%; Pred. No. 3.3e-210;
Matches 929; Conservative 0; Mismatches 0; Indels 96; Gaps 1;

QY 1 ATGGAAAGTATGAAATAATTAGCTAAGACTGGAGAGGGTCTTATGGGGTGTATTCAA 60
Db 1 ATGGAAAGTATGAAATAATTAGCTAAGACTGGAGAGGGTCTTATGGGGTGTATTCAA 60
QY 61 TGCAGAAACAAACCTCTGGACAAGTAGTAGCTGTTAAATAATTTGTGGAATCTGAAGAT 120
Db 61 TGCAGAAACAAACCTCTGGACAAGTAGTAGCTGTTAAATAATTTGTGGAATCTGAAGAT 120
QY 121 GATCCTGTGTAAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACACAT 180
Db 121 GATCCTGTGTAAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACACAT 180
QY 181 CCAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAGGAAATGCAATTTAGTTT 240
Db 181 CCAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAGGAAATGCAATTTAGTTT 240
QY 241 GAATACGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGGAGTGTCTGAT 300
Db 241 GAATACGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGGAGTGTCTGAT 300
QY 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTATATACAT 360
Db 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTATATACAT 360
QY 361 AACTGTATTCACAGAGATATAAAACCTGAAATATTTCTAATAACTAAGCAAGGAATAATC 420
Db 361 AACTGTATTCACAGAGATATAAAACCTGAAATATTTCTAATAACTAAGCAAGGAATAATC 420
QY 421 AAGATTGTGACTTCGGGTTTGCACAAATTTCTGATTCAGGAGATGCCTACACCGATTAT 480
Db 421 AAGATTGTGACTTCGGGTTTGCACAAATTTCTGATTCAGGAGATGCCTACACCGATTAT 480

QY 481 GTAGCTACGAGATGGTACCGAGCTCCTGAACCTTCTTGTGGAGATACCTCAGTATGGTTCT 540
Db 481 GTAGCTACGAGATGGTACCGAGCTCCTGAACCTTCTTGTGGAGATACCTCAGTATGGTTCT 540
QY 541 TCAGTCGATATATGGGCTATTGGTTGTGTTTTTGCAGAGCTCCTGACAGGCCAGCCACTG 600
Db 541 TCAGTCGATATATGGGCTATTGGTTGTGTTTTTGCAGAGCTCCTGACAGGCCAGCCACTG 600
QY 601 TGGCCTGGAATCAGATGTGGACCAACTTTATCTGATAATCAGAAACACTAGTAGAGACG 660
Db 601 TGGCCTGGAATCAGATGTGGACCAACTTTATCTGATAATCAGAAACACT----- 650
QY 661 GGGTTTCGCCATGTTGACCCAGGCTGGTCTCGAACTCTTGACGTCAAGTGATCCACCTGCC 720
Db 651 ----- 650
QY 721 GTAGCCTCTCAAGTGTGGAATTACAGGAAATTAATCCCAAGACATCAATCAATCTTT 780
Db 651 -----AGGAAATTAATCCCAAGACATCAATCAATCTTT 684
QY 781 AAAAGTAACGGTTTTTCCATGGCATCAGTATACCTGAGCCAGACATGGAAACTCTT 840
Db 685 AAAAGTAACGGTTTTTCCATGGCATCAGTATACCTGAGCCAGACATGGAAACTCTT 744
QY 841 GAGGAAAGTTCTCAGATGTTTCATCCTGTGGCTCTGAACTTCATGAAGGGGTGCTGAAG 900
Db 745 GAGGAAAGTTCTCAGATGTTTCATCCTGTGGCTCTGAACTTCATGAAGGGGTGCTGAAG 804
QY 901 ATGAATCCAGATGACAGATTAAACCTGTTCCTCCAACTCCTGGAGAGCTCCTACTTTGATTCT 960
Db 805 ATGAATCCAGATGACAGATTAAACCTGTTCCTCCAACTCCTGGAGAGCTCCTACTTTGATTCT 864
QY 961 TTTCAAGAGGCCCAATTAAGAAAGAAAGACCGTAATGAAGGAAGAAACAGAGACGCCAA 1020
Db 865 TTTCAAGAGGCCCAATTAAGAAAGAAAGACCGTAATGAAGGAAGAAACAGAGACGCCAA 924
QY 1021 CAGGT 1025
Db 925 CAGGT 929
RESULT 5
AAD30557
ID AAD30557 standard; cDNA; 1790 BP.
XX
AC AAD30557;
XX
DT 21-MAY-2002 (first entry)
XX
DE Humar. kinase polypeptide (PKIN-10) cDNA.
XX
KW Human; kinase polypeptide; PKIN-10; gene therapy; Addison's disease;
KW leukaemia; immune disorder; lymphoma; melanoma; developmental disorder;
KW acquired immunodeficiency syndrome; AIDS; allergy; anaemia; hypertension;
KW asthma; Crohn's disease; rheumatoid arthritis; bursitis; atherosclerosis;
KW cirrhosis; hepatitis; psoriasis; Cushing's syndrome; cytostatic; cancer;
KW cholestasis; cardiac; cardiovascular disorder; Niemann-Pick's disease;
KW lipid disorder; fatty liver; Gaucher's disease; myocardial infarction;
KW drug screening; transgenic animal; antiinflammatory; hepatotropic;
KW hypotensive; anti-HIV; enzyme; ss.
XX
OS Homo sapiens.
XX
XX
FH Key Location/Qualifiers
CDS 239..1267
FT /*tag= a
FT /product= "Human PKIN-10"
XX
PN WO200208399-A2.
XX
PD 31-JAN-2002.
XX

KW endometriosis; endocrine disorder; hyperparathyroidism; hyperlipidemia;
KW hypercholesterolaemia; hypoglycaemia; obesity; Reiter's syndrome;
KW connective tissue disorder; osteoporosis; infection.

XX Homo sapiens.
XX WO200279473-A2.
XX
XX 10-OCT-2002.
XX
XX 09-JAN-2002; 2002WO-US01009.
XX
XX 12-JAN-2001; 2001US-261622P.
XX 16-JAN-2001; 2001US-261864P.
XX 16-JAN-2001; 2001US-261865P.
XX 17-JAN-2001; 2001US-262164P.
XX 17-JAN-2001; 2001US-262207P.
XX 17-JAN-2001; 2001US-262208P.
XX 17-JAN-2001; 2001US-262209P.
XX 17-JAN-2001; 2001US-262215P.
XX 18-JAN-2001; 2001US-263102P.
XX 19-JAN-2001; 2001US-262599P.
XX 19-JAN-2001; 2001US-263063P.
XX 19-JAN-2001; 2001US-263064P.
XX 19-JAN-2001; 2001US-263065P.
XX 19-JAN-2001; 2001US-263069P.
XX 19-JAN-2001; 2001US-263077P.
XX 19-JAN-2001; 2001US-263329P.
XX 19-JAN-2001; 2001US-263330P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Panzer SR, Lincoln SE, Altus CM, Dufour GE, Hillman JL, Jones AL;
XX Dam TC, Liu TF, Harris B, Flores V, Daffo A, Marwaha R, Chen AJ;
XX Chang SC, Gerstin EH, Peralta CH, David MH, Lewis SA;
XX W22; 2003-040680/03.
XX P-PSDB; ABU05290.
XX
XX New human diagnostic and therapeutic (DITHP) polynucleotides and
XX polypeptides, useful for diagnosing, preventing or treating diseases,
XX e.g. cancer, AIDS, Parkinson's disease, or autoimmune/inflammatory
XX diseases -
XX
XX Claim 1; Page 267; 331pp; English.
XX
XX This invention relates to the cDNA and protein sequences of fifty six
XX polynucleotides for diagnostics and therapeutics (DITHP). The human
XX DITHP polynucleotides and polypeptides are useful for diagnosing,
XX preventing or treating diseases associated with, as well as effects of
XX exogenous compounds, on the expression of human molecules, such as cell
XX proliferative diseases (e.g. cancer, arteriosclerosis atherosclerosis,
XX psoriasis, primary thrombocytopenia), autoimmune/inflammatory diseases
XX (e.g. anaemia, asthma, autoimmune thyroiditis, Crohn's disease, diabetes
XX mellitus, glomerulonephritis, gout, multiple sclerosis, rheumatoid
XX arthritis, uveitis, acquired immunodeficiency disease; AIDS; allergies;
XX neurological disorders (e.g. stroke, Alzheimer's disease, dementia,
XX mental retardation, Parkinson's disease, epilepsy), gastrointestinal
XX (e.g. ulcer, cirrhosis), reproductive (e.g. infertility, endometriosis),
XX endocrine disorders (e.g. hyperparathyroidism, hypercholesterolaemia,
XX hypoglycaemia, hyperlipidemia, obesity), connective tissue disorders
XX (e.g. osteoporosis, Reiter's syndrome), or infections (e.g. bacterial,
XX viral, fungal, parasitic, protozoal). The DITHP sequences may be used to
XX generate hybridisation probes useful in chromosomal mapping of naturally
XX occurring genomic sequences. They are also useful in designing probes
XX useful in diagnostic assays. The polynucleotides are useful as molecular
XX weight markers, or as antigen to elicit an immune response. The
XX present sequence represents a human diagnostics and therapeutics (DITHP)
XX cDNA sequence of the invention.

SQ Sequence 1678 BP; 541 A; 308 C; 336 G; 493 T; 0 other;
Query Match 73.3%; Score 762.8; DB 25; Length 1678;

Best Local Similarity 90.0%; Pred. No. 5.2e-194;		Matches 925; Conservative 0; Mismatches 2; Indels 101; Gaps 6;	
Qy	1	ATGGAAAAGTATGAAAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGGTTGTATTCAA	60
Db	131	ATGGAGAAGTATGAAAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGGTTGTATTCAA	190
Qy	61	TGCAGA-AACAAAACCTCTGGACAAGTAGTAGCTGTAAAAAATTTGTGGAATCTGAAGA	119
Db	191	TGCAGATAACAAAACCTCTGGACAAGTAGTAGCTGTAAAAAATTTGTGGAATCTGAAGA	250
Qy	120	TGATCCTGTTGTTAAGAAAATAGCACTAAGAGAAAATACGTA-TGTTGAAGCAATTAAC	178
Db	251	TGATCCTGTTGTTAAGAAAATAGCACTAAGAGAAAATACGTAAGCAATTAAC	310
Qy	179	ATCCAAATCTTGTAACCTCATCGAGGTGTTTCAGGAGAAAAGGAAAATGCAATTAGTTT	238
Db	311	ATCCAAATCTTGTAACCTCATCGAGGTGTTTCAGGAGAAAAGGAAAATGCAATTAGTTT	370
Qy	239	TTGAATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAACCCAAATGGAGTTGCTG	298
Db	371	TTGAATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAACCCAAATGGAGTTGCTG	430
Qy	299	ATGGAGTGATCAAAAGCGTATTATGSCAAACACTTCAAGCTCTTAATTTCTGTCATATAC	358
Db	431	ATGGAGTGATCAAAAGCGTATTATGSCAAACACTTCAAGCTCTTAATTTCTGTCATATAC	490
Qy	359	ATAACTGTATTCACAGAGATATAAAACCTGAAAATATTCTTAATAAAGCAAGGAATAA	418
Db	491	ATAACTGTATTCACAGAGATATAAAACCTGAAAATATTCTTAATAAAGCAAGGAATAA	550
Qy	419	TCAAGATTTGTGACTTCGGGTTTGCAACAATTTCTGATTCAGAGATGCCTACACCGATT	478
Db	551	TCAAGATTTGTGACTTCGGGTTTGCAACAATTTCTGATTCAGAGATGCCTACACCGATT	610
Qy	479	ATGTAGCTA-CSAGATGGTACCGAGCT-CCTGAACCTTCTTGTGGGAGATACT-CAGTATG	535
Db	611	ATGTAGCTAGCGAGATGGTACCGAGCTCCCTGAACTTCTTGTGGGAGATACTCCAGTATG	670
Qy	536	GTTCTTCAGTCGATATATGGGCTATTGGTTGTGTTTTTGCAGAGCTCCTGACAGGCCAGC	595
Db	671	GTTCTTCAGTCGATATATGGGCTATTGGTTGTGTTTTTGCAGAGCTCCTGACAGGCCAGC	730
Qy	596	CACGTGTGGCCTGGAATAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGTAG	655
Db	731	CACGTGTGGCCTGGAATAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACT	785
Qy	656	AGACGGGTTTCGCCATGTTTGACCGGCTGGTCTCGAACTCTTGACGTCGAGTCCAC	715
Db	786	AGACGGGTTTCGCCATGTTTGACCGGCTGGTCTCGAACTCTTGACGTCGAGTCCAC	785
Qy	716	CTGCCGTAGCCTCTCAAAGTGCTGGAATTACAGGAAAATTAATCCCAAGACATCAATCAA	775
Db	786	CTGCCGTAGCCTCTCAAAGTGCTGGAATTACAGGAAAATTAATCCCAAGACATCAATCAA	814
Qy	776	TCTTTAAAAGTAACGGGTTTTTCCATGGCATCAGTATACCTGAGCCAGAGACATGGAAA	835
Db	815	TCTTTAAAAGTAACGGGTTTTTCCATGGCATCAGTATACCTGAGCCAGAGACATGGAAA	874
Qy	836	CTCTTGAGGAAAAGTTCTCAGATGTTTCATCCTGTGGCTCTGAACCTTCATGAAGGGTGTG	895
Db	875	CTCTTGAGGAAAAGTTCTCAGATGTTTCATCCTGTGGCTCTGAACCTTCATGAAGGGTGTG	934
Qy	896	TGAAGATGAATCCAGATGACAGATTAACCTGTTCCCAACTCTCTGGAGAGCTCCTACTTTG	955
Db	935	TGAAGATGAATCCAGATGACAGATTAACCTGTTCCCAACTCTCTGGAGAGCTCCTACTTTG	994
Qy	956	ATTCTTTTCAAGAGGCCCAAAATTAAGAAAAGCAAGCTTAATGAAGGAAAGAAACAGAGAC	1015
Db	995	ATTCTTTTCAAGAGGCCCAAAATTAAGAAAAGCAAGCTTAATGAAGGAAAGAAACAGAGAC	1054
Qy	1016	GCCAAACAG 1023	

Db 1055 GCCAACAG 1062

RESULT 7
AAS06725
ID AAS06725 standard; cDNA; 1093 BP.
XX
AC AAS06725;
XX
DT 12-SEP-2001 (first entry)
DE Po-ynucleotide sequence encoding human protein kinase #25.
XX
KW Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
KW metabolic disorder; immune related disease; neurological disorder;
KW neurodegenerative disorder; inflammatory disorder; infectious disease;
KW reproductive disorder; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200138503-A2.
XX
PD 31-MAY-2001.
XX
PF 22-NOV-2000; 2000WO-US32085.
XX
PR 24-NOV-1999; 99US-0167482.
XX
PA (SUGE-) SJGEN INC.
XX
PI Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
PI Flanagan P, Ciary D;
XX
DR WPI; 2001-343950/36.
DR P-PSDB; AAU03525.
XX
PT Nucleic acids encoding human kinase polypeptides, useful for preventing
PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and
PT neuronal-associated diseases, and microbial infections -
XX
PS Example 1; Figure 1; 433pp; English.
XX
CC AAS06701-AAS06757 encode for novel human protein kinases #1-57. The
CC novel protein kinases have been identified as members of the tyrosine
CC or serine/threonine kinase (PTK and STK) families. The polynucleotides
CC encoding protein kinases and the polypeptides may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate kinase expression. For example, they may be used to treat
CC cancers (especially cancers of haematopoietic origin), cardiovascular
CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
CC immune related diseases (e.g. rheumatoid arthritis), neurological
CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).
CC Additionally, polynucleotides encoding protein kinases may be
CC used for gene therapy and as DNA probes in diagnostic assays.
CC The protein kinase polypeptides may be used as antigens in the production
CC of antibodies against the protein kinases and in assays to identify
CC modulators of protein kinase expression and activity.
XX
SQ Sequence 1083 BP; 366 A; 199 C; 225 G; 293 T; 0 other;

Query Match 72.7%; Score 757; DB 22; Length 1083;
Best Local Similarity 86.1%; Pred. No. 1.6e-192;
Matches 927; Conservative 0; Mismatches 0; Indels 150; Gaps 2;

Qy 1 ATGGAAAAGTATGAAAATTAGCTAAGACTGGAGAAGGGTCTTATGGGTGTGATTCAAA 60
Db 1 ATGGAAAAGTATGAAAATTAGCTAAGACTGGAGAAGGGTCTTATGGGTGTGATTCAAA 60
Qy 61 TGCAGAAACAAAACCTCTGGACAAGTAGTAGTGTGTTAAAAAATTTGTGGAATCTGAAGAT 120
Db 61 TGCAGAAACAAAACCTCTGGACAAGTAGTAGTGTGTTAAAAAATTTGTGGAATCTGAAGAT 120

QY 121 GATCCTGTTGTTAAGAAAATAGCACTAAGAGAAAATACGTAATGTTGAAGCAATTTAAACAT 180
Db 121 GATCCTGTTGTTAAGAAAATAGCACTAAGAGAAAATACGTAATGTTGAAGCAATTTAAACAT 180
QY 181 CCAAAATCTGTGAACCTCATCGAGGTGTTCAAGGAGAAAAGGAAAATGCAATTTAGTTTTT 240
Db 181 CCAAAATCTGTGAACCTCATCGAGGTGTTCAAGGAGAAAAGGAAAATGCAATTTAGTTTTT 240
QY 241 GAATACCTGTGATCATACACTTTTAAATGAGCTGGAAGAAAACCCCAATGGAGTTGCTGAT 300
Db 241 GAATACCTGTGATCATACACTTTTAAATGAGCTGGAAGAAAACCCCAATGGAGTTGCTGAT 300
QY 301 GGAGTGATCAAAAAGCGTATTATGCAAAACACTTCAAGCTCTTAATTTCTGTCAATATACAT 360
Db 301 GGAGTGATCAAAAAGCGTATTATGCAAAACACTTCAAGCTCTTAATTTCTGTCAATATACAT 360
QY 361 AACTGTATTACAGAGATATAAAACCTGAAAATATTCTAATAACTAAGCAAGGAATAATC 420
Db 361 AACTGTATTACAGAGATATAAAACCTGAAAATATTCTAATAACTAAGCAAGGAATAATC 420
QY 421 AAGATTTGTGACTTCGGGTTTGCAAAAATTCGTA----- 454
Db 421 AAGATTTGTGACTTCGGGTTTGCAAAAATTCGTAATTCACATGTTGGGAGAAATGAC 480
QY 455 -----TTCCAGGAGATGCTTACACCGATTATGTAGCT 486
Db 481 TTACCTGATCTGATTGATGCAATTCAGAGTTCAGGAGATGCTTACACCGATTATGTAGCT 540
QY 487 ACAGATGGTACCGAGCTCCTGAACCTTCTTGTGGGAGATACACTAGTATGGTCTTCAGTC 546
Db 541 ACAGATGGTACCGAGCTCCTGAACCTTCTTGTGGGAGATACACTAGTATGGTCTTCAGTC 600
QY 547 GATATATGGGCTATTGGTTGTGTTTTCAGAGAGCTCCTGACAGGCCAGCCACTGTGGCCT 606
Db 601 GATATATGGGCTATTGGTTGTGTTTTCAGAGAGCTCCTGACAGGCCAGCCACTGTGGCCT 660
QY 607 GGAATAATCAGATGTGGACCAACTTATCTGATAATCAGAACACTAGTAGAGACGGGGTTT 666
Db 661 GGAATAATCAGATGTGGACCAACTTATCTGTGATAATCAGAACACT----- 704
QY 667 CGCCATGTTGACACAGGCTGGTCTCGAACTCTTGACGTCAGTATCCACCTGCCGTAGCC 726
Db 705 ----- 704
QY 727 TCTCAAAGTGTGGAATTACAGGAAAATTAATCCCAAGACATCAATCAATCTTTAAAGT 786
Db 705 -----AGGAAAATTAATCCCAAGACATCAATCAATCTTTAAAGT 744
QY 787 AACGGGTTTTCATGTCATGTCGCTCTGAACCTTCAAGAGAGGAGTCTGAAGATGAAT 846
Db 745 AACGGGTTTTCATGTCATGTCGCTCTGAACCTTCAAGAGAGGAGTCTGAAGATGAAT 804
QY 847 AAGTTCTCAGATGTTCACTGTCGCTCTGAACCTTCAAGAGAGGAGTCTGAAGATGAAT 906
Db 805 AAGTTCTCAGATGTTCACTGTCGCTCTGAACCTTCAAGAGAGGAGTCTGAAGATGAAT 864
QY 907 CCAGATGACAGATTAACTGTTCCCAACTCCTGGAGAGCTCCTACTTTGATTTCTTTCAA 966
Db 865 CCAGATGACAGATTAACTGTTCCCAACTCCTGGAGAGCTCCTACTTTGATTTCTTTCAA 924
QY 967 GAGGCCCAATTAAGAAAAGCACTAATGAAGGAAGAAAACAGACGACGCAACAG 1023
Db 925 GAGGCCCAATTAAGAAAAGCACTAATGAAGGAAGAAAACAGACGACGCAACAG 981

RESULT 8
AAD03818
ID AAD03818 standard; cDNA; 1819 BP.
XX
AC AAD03818;
XX
DT 19-JUN-2001 (first entry)

XX	
DE	Human kinase cDNA #7.
XX	
XW	Human; kinase; gene therapy; bioreactor; mental disorder;
XW	biological disorder; polymorphism; ss.
XX	
OS	Homo sapiens.

Key	Location/Qualifiers
5'UTR	1..137
FT	/*tag= a
F+	138..701
FT	/*tag= b
FT	/product= "Human kinase #1"
FT	702..1819
F+	/*tag= c
misc_difference	1684
FT	/*tag= d
FT	/note= "polymorphism site;"
FT	

PW0200123579-A1.
XX
XX PD C5-APR-2001.
XX
PF 27-SEP-2000; 2000NO-US26621.
XX
PR 28-SEP-1999; 99US-0156511.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
XX
DR WPI; 2001-266166/27.
DR P-PSDB; AAE00490.
XX
PT New isolated human kinase polynucleotide useful for generating
PT antibodies, as reagents in diagnostic assays and for screening for
PT compounds useful for treating mental, biological or medical diseases -
XX
PS Disclosure; Page 34-35; 38pp; English.

The present sequence encodes a novel human protein (NHP) which is a human kinase. A polymorphism was identified in the 3' UTR of the present sequence. The human kinases share structural similarity with animal kinases, more particularly serine or threonine protein kinases. Human kinase cDNA is useful for the detection of mutant human kinase for the diagnosis of disease, and also as a therapeutic. It is useful for screening drugs effective in the treatment of symptomatic or phenotypic manifestations perturbing the normal function of NHP in the body. The NHP nucleotide sequences are useful for generation of antibodies, as reagents in diagnostic assays, for the identification of other cellular gene products related to human kinases, and as reagents in assays for screening compounds that are useful for treating mental, biological or medical disorders. NHP oligonucleotides are used as probes. The labelled NHP probes are useful for screening human genomic library for identifying polymorphisms and as primers in amplification assays to detect mutations within the exons, introns and splice sites that can be used in diagnostics and pharmacogenomics. Nucleotide constructs encoding NHP products are used to genetically engineer cells in vivo that functions as bioreactors in the body delivering a continuous supply of NHP to the body. Nucleotide constructs encoding functional NHPs are used in gene therapy for the modulation of NHP expression.

Sequence 1819 BP; 539 A; 368 C; 406 G; 505 T; 1 other;

Query Match	68.1%	Score 709;	DB 22;	Length 1819;
Best Local Similarity	77.0%	Pred. No. 1.4e-179;		
Matches 1025;	Conservative	0;	Mismatches 0;	Indels 306;
				Gaps 1;

QY 1 ATGGAAAGTATGAAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGGTTGTATTCTAAA 60

Db	138	ATGGAAAAGTATGAAAAAATTAGCTAAGACTGGAGAAAGGGTCTTATGGGGTTGTATTCAA	197
QY	61	TGCAGAAACAAAACCTCTCGACAAGTAGTAGCTGTTAAAAAAATTTGTGGAATCTGAAGAT	120
Db	198	TGCAGAAACAAAACCTCTCGACAAGTAGTAGCTGTTAAAAAAATTTGTGGAATCTGAAGAT	257
QY	121	GATCCTGTTGTTAAGAAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAATAACAT	180
Db	258	GATCCTGTTGTTAAGAAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAATAACAT	317
QY	181	CCAAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAAGGAAAATGCAATTTAGTTTTT	240
Db	318	CCAAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAAGGAAAATGCAATTTAGTTTTT	377
QY	241	GAATACTGTGATCATACACTTTTAAATGAGCTGGAAGAAAACCCAAATGGAGTTGCTGAT	300
Db	378	GAATACTGTGATCATACACTTTTAAATGAGCTGGAAGAAAACCCAAATGGAGTTGCTGAT	437
QY	301	GGAGTGATCAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGTCTATATACAT	360
Db	438	GGAGTGATCAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGTCTATATACAT	497
QY	361	AACCTGATTACAGAGATATAAAACCTGAAATATTCTAATAACTAAGCAAGGAATAATC	420
Db	498	AACCTGATTACAGAGATATAAAACCTGAAATATTCTAATAACTAAGCAAGGAATAATC	557
QY	421	AAGATTTGTGACTTCGGGTTTGACAAAATTCG-	453
Db	558	AAGATTTGTGACTTCGGGTTTGACAAAATTCGAGTTGAGCTTCAATCTTCTCTGGTGCC	617
QY	454	-	453
Db	618	TCCTTGATTTGGCTTAANTAGTTGACCTTCTGAATTTCTTTCTGCCAATTCAGAGATTTTT	677
QY	454	-	453
Db	678	CTCCTGGCTTGGATCCATTGCTGACACAGTGTTTCCACCATGGGGCCAGGCTCATCTCGA	737
QY	454	-	453
Db	738	ACTTCTGGCCTCAAGTGATCCTTCCACCTCGGCTCCCAAAGTGCTGGATTGCAAGTGCTG	797
QY	454	-	453
Db	798	AGCCACCGTGGCCAGCCAGATTTTTCAAACAATAACTACTGAGAGCTCAACAAGATTGTTT	857
QY	454	-	453
Db	858	TTAGTGGAAACACAAATTCGAACAAATTTCTTGAGAACGCATTCAGAGAGATGCCTACACC	917
QY	475	GATTATGTAGCTACGAGATGGTACCGAGCTCCTGAACTTCTTGTGGGAGATACTCAGTAT	534
Db	918	GATTATGTAGCTACGAGATGGTACCGAGCTCCTGAACTTCTTGTGGGAGATACTCAGTAT	977
QY	535	GGTCTTCAGTCGATATATGGGCTATTGGTTGTTTTCAGAGCTCCTGACAGGCCAG	594
Db	978	GGTCTTCAGTCGATATATGGGCTATTGGTTGTTTTCAGAGCTCCTGACAGGCCAG	1037
QY	595	CCACTGTGGCCTGGAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGTA	654
Db	1038	CCACTGTGGCCTGGAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGTA	1097
QY	655	GAGACGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTCTTGACGTCAGTGATCCA	714
Db	1098	GAGACGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTCTTGACGTCAGTGATCCA	1157
QY	715	CCTGCCCTAGCCTCTCAAAGTGCTGGAATTACAGGAAAATTAATCCCAAGACATCAATCA	774
Db	1158	CCTGCCCTAGCCTCTCAAAGTGCTGGAATTACAGGAAAATTAATCCCAAGACATCAATCA	1217
QY	775	ATCTTTAAAAAGTAACCGGTTTTTCCATGGCATCAGTATACCTGAGCCAGAACATGGAA	834

Db 1218 ATCTTTAAAGTAACGGGTTTTTCATGGCATCAGTATACCTGAGCCAGAGACATGGAA 1277
QY 835 ACTCTTGAGGAAAGTTCTCAGATGTTTCATCGTGTGGCTCTGAACCTCATGAAGGGGTGT 894
Db 1278 ACTCTTGAGGAAAGTTCTCAGATGTTTCATCGTGTGGCTCTGAACCTCATGAAGGGGTGT 1337
QY 895 CTGAAGATGAATCCAGATGACAGATTAAACCTGTTCCCAACTCCTGAGAGCTCCTACTTT 954
Db 1338 CTGAAGATGAATCCAGATGACAGATTAAACCTGTTCCCAACTCCTGAGAGCTCCTACTTT 1397
QY 955 GATTCTTTCAAGAGGCCCCAAATTAAAGAAAGACCGTAATGAAGGAAGAAACAGAGA 1014
Db 1398 GATTCTTTCAAGAGGCCCCAAATTAAAGAAAGACCGTAATGAAGGAAGAAACAGAGA 1457
QY 1015 CGCCACACAGGT 1025
Db 1458 CGCCACACAGGT 1468

RESULT 9
AAI64248
ID AAI64248 standard; cDNA; 882 BP.
XX

AC AAI64248;

XX 08-MAR-2002 (first entry);

DT Human kinase 14257 cDNA.

DE Human kinase 14257 cDNA.

XX

KW Protein kinase; enzyme; cytostatic; osteopathic; hepatotropic;
KW antidiabetic; neuroprotective; antiarthritic; dermatological;
KW immunosuppressive; antiinflammatory; antithyroid; antipsoriatic;
KW ophthalmological; antiallergic; antiasthmatic; antiatherosclerotic;
KW hypotensive; vasotropic; antiarrhythmic; virucide; anorectic;
KW metabolic; immunomodulator; analgesic; cellular proliferative disorder;
KW cancer; acute lymphoblastic leukaemia; Hodgkin's disease;
KW bone metabolism disorder; osteoporosis; immune system disorder;
KW inflammatory; diabetes mellitus; osteoarthritis; asthma;
KW cardiovascular disorder; hypertension; coronary artery disease;
KW endothelial cell disorder; psoriasis; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FH 5'UTR 1..128
FT /*tag= a
FT 129..815
FT /*tag= b
FT /product= "protein kinase 14275"
FT 816..882
FT /*tag= c

XX WO200179488-A2.

XX 25-OCT-2001.

XX 13-APR-2001; 2001WO-US12188.

XX 13-APR-2000; 2000US-196910P.

XX (MILL-) MILLENNIUM PHARM INC.

XX Kapeller-libermann R;

XX WPI; 2002-034355/04.

XX P-PSDB; AAG78547.

XX New 14257 polypeptides (protein kinases), useful as diagnostic targets
PT and therapeutic agents for controlling cellular proliferative and/or
PT differentiative disorder, bone disorders, immune disorders and
PT cardiovascular disorders -

XX Claim 2; Fig 1A; 98pp; English.

PS

XX The invention relates to an isolated 14257 polypeptide and nucleic
CC acid encoding it. The 14257 protein is a protein kinase that acts as a
CC modulating agent in regulating a variety of cellular processes,
CC including cell proliferation, differentiation, growth and division.
CC The activity of the protein of the invention may be described as;
CC cytostatic; osteopathic; hepatotropic; antidiabetic; neuroprotective;
CC antiarthritic; dermatological; immunosuppressive; antiinflammatory;
CC antithyroid; antipsoriatic; ophthalmological; antiallergic;
CC antiasthmatic; antiatherosclerotic; hypotensive; vasotropic;
CC antiarrhythmic; virucide; anorectic; metabolic; immunomodulator and
CC analgesic. The protein of the invention may act as a novel diagnostic
CC target or therapeutic agent controlling certain disorders, for example
CC kinase-associated or other 14257-associated disorders. These may include
CC cellular proliferative disorders such as cancers e.g. acute lymphoblastic
CC leukemia or Hodgkin's disease. Other disorders include bone metabolism
CC disorders such as osteoporosis, disorders of the immune system, e.g.
CC inflammatory, diabetes mellitus, osteoarthritis and asthma. Proteins of
CC the invention may also be of use as therapeutic agents in cardiovascular
CC disorders such as hypertension and coronary artery disease, and some
CC endothelial cell disorders, including psoriasis. The current
CC sequence represents a human kinase 14257 cDNA.

XX
SQ Sequence 882 BP; 288 A; 150 C; 198 G; 245 T; 1 other;

Query Match 61.9%; Score 644; DB 24; Length 882;
Best Local Similarity 99.2%; Pred. No. 2.6e-162;
Matches 647; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ATGGAAAGTATGAAATAATTAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAA 60
Db 129 ATGGAAAGTATGAAATAATTAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAA 188
QY 61 TGCAGAAACAAACCTCTGGACAAGTAGTAGCTGTTAAAAAATTTGTGGAATCTGAAGAT 120
Db 189 TGCAGAAACAAACCTCTGGACABGTAGTAGCTGTTAAAAAATTTGTGGAATCTGAAGAT 248
QY 121 GATCCTGTTGTTAAGAAAATAGCACTAAGAGAAAATACGTTATCTTGAAGCAATTTAAACAT 180
Db 249 GATCCTGTTGTTAAGAAAATAGCACTAAGAGAAAATACGTTATCTTGAAGCAATTTAAACAT 308
QY 181 CCAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAGGAAAATGCATTTAGTTT 240
Db 309 CCAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAGGAAAATGCATTTAGTTT 368
QY 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGGAGTTGCTGAT 300
Db 369 GAATACTGTGATCATGCACTTTTAAATGAGCTGGAAGAAACCCAAATGGAGTTGCTGAT 428
QY 301 GGAGTGATCAAAAGCGTATTATGGCAACACCTTCAAGCTCTTAATTTCTGCATATACAT 360
Db 429 GGAGTGATCAAAAGCGTATTATGGCAACACCTTCAAGCTCTTAATTTCTGCATATACAT 488
QY 361 AACTGTATTCACAGAGATATAAAACCTGAAAATATTCTAATAACTAAGCAAGGAATAATC 420
Db 489 AACTGTATTCACAGAGATGTAAACCTGAAAATATTCTAATAACTAAGCAAGGAATAATC 548
QY 421 AAGATTTGTGACTTCGGGTTTGGCAAAATTTCTGATTCAGAGAGATGCCTACACCGATTAT 480
Db 549 AAGATTTGTGACTTCGGGTTTGGCAAAATTTCTGATTCAGAGAGATGCCTACACCGATTAT 608
QY 481 GTAGCTACGAGATGGTACCGAGCTCCTGAACTTCTTTGTGGAGATACCTCAGTATGTTCT 540
Db 609 GTAGCTACGAGATGGTACCGAGCTCCTGAACTTCTTTGTGGAGATACCTCAGTATGTTCT 668
QY 541 TCAGTCGATATATGGGCTATTGGTTGTGTTTTTTCAGAGAGCTCCTGACAGGCCAGCCACTG 600
Db 669 TCAGTCGATATATGGGCTATTGGTTGTGTTTTTTCAGAGAGCTCCTGACAGGCCAGCCACTG 728
QY 601 TGGCCTGGAATAATCAGATGTGGACCAACTTTTATCTGATAATCAGAACACTAG 652
Db 729 TGGCCTGGAATAATCAGATGTGGACCAACTTTTATCTGATAATCAGAACACTAG 780

RESULT 10
ABX34679
ID ABX34679 standard; cDNA; 1281 BP.
XX AC ABX34679;
XX DT 13-FEB-2003 (first entry)
XX DE Human mddt cDNA SEQ ID 240.
XX KW MDDT; human; disease detection and treatment molecule polypeptide;
KW anti-inflammatory; immunosuppressive; osteopathic; cytostatic; anti-HIV;
KW haemostatic; nephrotropic; antianaemic; antipsoriatic; hepatotropic;
KW gene therapy; protein replacement therapy; cell proliferative disorder;
KW cancer; adenocarcinoma; leukaemia; lymphoma; melanoma; sarcoma;
KW anaemia; Crohn's disease; acquired immunodeficiency syndrome; AIDS;
KW Goodpasture's syndrome; inflammation; osteoporosis; thrombocytopaenia;
KW psoriasis; hepatitis; gene; ss.
XX OS Homo sapiens.
XX PN WO200279449-A2.
XX PD 10-OCT-2002.
XX PF 27-MAR-2002; 2002WO-US09944.
XX PR 28-MAR-2001; 2001US-279619P.
PR 29-MAR-2001; 2001US-280067P.
PR 29-MAR-2001; 2001US-280068P.
PR 16-MAY-2001; 2001US-291280P.
PR 17-MAY-2001; 2001US-291829P.
PR 17-MAY-2001; 2001US-291849P.
PR 19-JUN-2001; 2001US-299428P.
PR 20-JUN-2001; 2001US-299776P.
PR 20-JUN-2001; 2001US-300001P.
XX PA (INCY-) INCYTE GENOMICS INC.
XX PI Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
PI Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshey SR;
PI Daugherty SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstein EH;
PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
PI Flores V, Marwaha R, Lo A, Lan RY, Jrashka NE;
XX WP: 2003-058431/05.
DR P-PSDB; ABU11689.
XX PT New purified disease detection and treatment molecule proteins and
PT polynucleotides, useful for diagnosing, treating or preventing cancers
PT (e.g. leukemia or sarcoma), anemia, Crohn's disease, AIDS, osteoporosis
PT or hepatitis -
XX PS Claim 1; SEQ ID NO 240; 339pp - Sequence Listing; English.
XX CC This invention describes a novel disease detection and treatment molecule
CC polypeptide (MDDT) which has anti-inflammatory, immunosuppressive,
CC osteopathic, cytostatic, anti-HIV, haemostatic, nephrotropic,
CC antianaemic, antipsoriatic and hepatotropic activity. The polynucleotides
CC and the polypeptides of the invention can be used for gene therapy,
CC protein replacement therapy and are useful for treating a variety of
CC diseases or conditions. These polypeptides or polynucleotides are
CC particularly useful for diagnosing, treating or preventing cell
CC proliferative disorders (e.g. cancers including adenocarcinoma,
CC leukaemia, lymphoma, melanoma, myeloma or sarcoma), anaemia, Crohn's
CC disease, acquired immunodeficiency syndrome (AIDS), Goodpasture's
CC syndromes, inflammation, osteoporosis, thrombocytopaenia, psoriasis or
CC hepatitis. ABX34440-ABX34835 encode the MDDT polypeptides represented in
CC ABU11450-ABU11845, described in the disclosure of the invention.
CC NOTE: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 1281 BP; 428 A; 191 C; 317 G; 345 T; 0 other;
Query Match. 52.2%; Score 543.2; DB 25; Length 1281;
Best Local Similarity 99.5%; Pred. No. 2.8e-135;
Matches 545; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ATGGAAAGTATGAAAAATAGCTAAGACTGGAGAGGGTCTTATCGGGTTGATTCAAA 60
DB 704 ATGGAAAGTATGAAAAATAGCTAAGACTGGAGAGGGTCTTATCGGGTTGATTCAAA 763
QY 61 TGCAGAAACAAACCTCTGGACAAGTAGTAGCTGTAAATAATTTGTGGAATCTGAAGAT 120
DB 764 TGCAGAAACAAACCTCTGGACAAGTAGTAGCTGTAAATAATTTGTGGAATCTGAAGAT 823
QY 121 GATCCTGTTGTTAAGAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAT 180
DB 824 GATCCTGTTGTTAAGAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAT 883
QY 181 CCAATCTTGTGACCTCATCGAGGTGTTCAAGAGAAAAGGAAAATGCATTGTTTTT 240
DB 884 CCAATCTTGTGACCTCATCGAGGTGTTCAAGAGAAAAGGAAAATGCATTGTTTTT 943
QY 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT 300
DB 944 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT 1003
QY 301 GGAGTGATCAAAAGCGTATTATGGCAAACACTTCAAGCTCTTAATTTCTGTCATATACAT 360
DB 1004 GGAGTGATCAAAAGCGTATTATGGCAAACACTTCAAGCTCTTAATTTCTGTCATATACAT 1063
QY 361 AACTGTATTACAGAGATATAAAACCTGAAATATTCTAATAACTAAGCAAGGAATAATC 420
DB 1064 AACTGTATTACAGAGATATAAAACCTGAAATATTCTAATAACTAAGCAAGGAATAATC 1123
QY 421 AAGATTGTGACTTCGGGTTTGACACAAATTTCCAGGAGATGCCATACACCGATTAT 480
DB 1124 AAGATTGTGACTTCGGGTTTGACACAAATTTCCAGGAGATGCCATACACCGATTAT 1183
QY 481 GTAGCTACGAGATGTTACCGAGCTCCTGAACTTCTTGTGGAGATACATCAGTATGTTCT 540
DB 1184 GCAGCTACGAGATGTTACCGAGCTCCTGAACTTCTTGTGGAGATACATCAGTATGTTCT 1243
QY 541 TCAGTCCA 548
DB 1244 TCAGTCAA 1251
RESULT 11
AAF44669
ID AAF44669 standard; cDNA; 911 BP.
XX AC AAF44669;
XX DT 27-MAR-2001 (first entry)
XX DE Novel protein kinase cDNA, SEQ ID NO: 49.
XX KW Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
KW immunosuppressive; cardiant; renal; antiinflammatory; antiasthmatic;
KW dermatological; antidiabetic; antiinfertility; gene therapy; vaccine;
KW immune disorder; cardiovascular disease; neurodegenerative disease;
KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;
KW inflammatory pelvic disease; multiple sclerosis; psoriasis; ss.
XX OS Homo sapiens.
XX PN WO200073469-A2.
XX PD 07-DEC-2000.
XX PF 26-MAY-2000; 2000WO-US14842.
XX

PR 28-MAY-1999; 99US-0136503.
XX (SUJE-) SUGEN INC.
PA Plowman GD, Martinez R, Whyte D, Sudersanam S;
XX WPI; 2001-032161/04.
DR P-PSDB; AAB65642.
XX Nucleic acids encoding kinase polypeptides, useful for diagnosing and
PT treating immune-related diseases and disorders, cardiovascular disease,
PT neurodegenerative diseases and/or cancers -
XX
PS Example 4; Fig 2; 310pp; English.
XX
CC The present sequence encodes a novel protein kinase. The nucleic acids
CC and the protein kinases they encode may be used in the treatment and
CC diagnosis of diseases associated with inappropriate kinase expression
CC such as immune-related diseases and disorders, cardiovascular disease,
CC neurodegenerative diseases and/or cancers. The nucleic acids and
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays. The kinase polypeptides may be used as antigens in the production
CC of antibodies of kinase expression and activity. Anti-kinase antibodies
CC and kinase antagonists may also be used to down regulate kinase
CC expression and activity. Diseases related to kinase expression and
CC activity include rheumatoid arthritis, atherosclerosis, autoimmune
CC disorders, complications of organ transplantation, myocardial infarction,
CC immune disorders, cardiomyopathies, strokes, renal failure,
CC oxidative-stress related disorders, chronic inflammatory bowel disease,
CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,
CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and
CC reproductive disorders.
XX
SQ Sequence 911 BP; 293 A; 182 C; 178 G; 258 T; 0 other;

Query Match 51.6%; Score 537; DB 22; Length 911;
Best Local Similarity 87.0%; Pred. No. 1.1e-133;
Matches 643; Conservative 0; Mismatches 0; Indels 96; Gaps 1;

QY 285 AAATGGAGTTGCTGATGGAGTGATCAAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAA 344
DB 1 AAATGGAGTTGCTGATGGAGTGATCAAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAA 60

QY 345 TTTCTGTCATATACATAACTGTATTTCACAGAGATATAAAACCTGAAATATTCTAATAAC 404
DB 61 TTTCTGTCATATACATAACTGTATTTCACAGAGATATAAAACCTGAAATATTCTAATAAC 120

QY 405 TAAGCAAGGAATANTCAAGATTTGTGACTTCGGGTTCGACAAATTCGTATCCAGGAGA 464
DB 121 TAAGCAAGGAATANTCAAGATTTGTGACTTCGGGTTCGACAAATTCGTATCCAGGAGA 180

QY 465 TGCCTACACCGGATTATGTAGCTACGAGATGGTACCGAGCTCCTGAACCTTCTTGGGAGA 524
DB 181 TGCCTACACCGGATTATGTAGCTACGAGATGGTACCGAGCTCCTGAACCTTCTTGGGAGA 240

QY 525 TACTCAGTATGGTTCTTCAGTCGATATATGGGCTATTGGTTGTGTTTTGACAGAGCTCCT 584
DB 241 TACTCAGTATGGTTCTTCAGTCGATATATGGGCTATTGGTTGTGTTTTGACAGAGCTCCT 300

QY 585 GACAGGCCAGCCACTGTGGCCTGGAAATCAGATGTGACCACTTTTATCTGATATATCAG 644
DB 301 GACAGGCCAGCCACTGTGGCCTGGAAATCAGATGTGACCACTTTTATCTGATATATCAG 360

QY 645 AACACTAGTAGAGACGGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTCTTGACGTC 704
DB 361 AACACT-----

QY 705 AAGTGATCCACCTGCCGTAGCCCTCTCAAAAGTCTGGAATTACAGGAAAAATTATCCCAAG 764
DB 367 -----AGGAAAAATTATCCCAAG 384

QY 765 ACATCAATCAATCTTTAAAGTAAACGGGTTTTTCCATGGCATCAGTATACCTGAGCCAGA 824
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

DB 385 ACATCAATCAATCTTTTAAAGTAACGGGTTTTTCCATGGCATCAGTATACCTGAGCCAGA 444
QY 825 AGACATGGAAACTCTTTGAGGAAAAAGTTCTCAGATGTTTCATCTGTGGCTCTGAACCTCAT 884
DB 445 AGACATGGAAACTCTTTGAGGAAAAAGTTCTCAGATGTTTCATCTGTGGCTCTGAACCTCAT 504
QY 885 GAAGGGGTGTCGAAGATGAATCCAGATGACAGATTAACCTGTTCCCAACTCCTGGAGAG 944
DB 505 GAAGGGGTGTCGAAGATGAATCCAGATGACAGATTAACCTGTTCCCAACTCCTGGAGAG 564
QY 945 CTCCTACTTTGATCTTTTCAAGAGGCCCAAAATTAAAGAAAAAGCAGGTAATGAAGGAAG 1004
DB 565 CTCCTACTTTGATCTTTTCAAGAGGCCCAAAATTAAAGAAAAAGCAGGTAATGAAGGAAG 624
QY 1005 AAACAGAAGACGCCAACAG 1023
DB 625 AAACAGAAGACGCCAACAG 643

RESULT 12
AAD03812
ID AAD03812 standard; cDNA; 561 BP.
XX
AC AAD03812;
XX 19-JUN-2001 (first entry)
DT Human kinase cDNA #1.
XX Human; kinase; gene therapy; bioreactor; mental disorder;
KW biological disorder; ss.
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..561
FT /*tag= a
FT /product= "Human kinase #1"
FT /note= "The coding region does not include stop codon"
FT /partial
XX
PK WO200123579-A1.
XX
PC 05-APR-2001.
XX
PF 27-SEP-2000; 2000WO-US26621.
XX
PR 28-SEP-1999; 99US-0156511.
XX (LEXI-) LEXICON GENETICS INC.
XX
PI Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
XX WPI; 2001-266166/27.
DR P-PSDB; AAE00490.
XX
XX New isolated human kinase polynucleotide useful for generating
PT antibodies, as reagents in diagnostic assays and for screening for
PT compounds useful for treating mental, biological or medical diseases -
XX
PS Disclosure; Page 27; 38pp; English.
XX
CC The present sequence is a cDNA encoding novel human protein
CC (NHP) known as human kinase. The human kinases share structural
CC similarity with animal kinases, more particularly serine or
CC threonine protein kinases. Human kinase cDNA is useful for the
CC detection of mutant human kinase for the diagnosis of disease,
CC and also as a therapeutic. It is useful for screening drugs
CC effective in the treatment of symptomatic or phenotypic
CC manifestations perturbing the normal function of NHP in the
CC body. The NHP nucleotide sequences are useful for generation of
CC antibodies, as reagents in diagnostic assays, for the
CC identification of other cellular gene products related to human

CC kinases, and as reagents in assays for screening compounds that
CC are useful for treating mental, biological or medical disorders.
CC NHP oligonucleotides are used as probes. The labelled NHP probes
CC are useful for screening human genomic library for identifying
CC polymorphisms and as primers in amplification assays to detect
CC mutations within the exons, introns and splice sites that can
CC be used in diagnostics and pharmacogenomics. Nucleotide construct
CC encoding NHP products are used to genetically engineer cells
CC in vivo that functions as bioreactors in the body delivering a
CC continuous supply of NHP to the body. Nucleotide constructs
CC encoding functional NHPs are used in gene therapy for the
CC modulation of NHP expression.

XX
SQ Sequence 561 BP; 189 A; 87 C; 112 G; 173 T; 0 other;

Query Match 43.7%; Score 454.4; DB 22; Length 561;
Best Local Similarity 99.8%; Pred. No. 1.2e-111;
Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGAAAGTATGAAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGGTTGTATTCAA 60
DB 1 ATGGAAAGTATGAAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGGTTGTATTCAA 60
QY 61 TGCAGAAACAAACCTCTGGACAAGTAGTAGCTGTTAAAAAATTTGTGGAATCTGAAGAT 120
DB 61 TGCAGAAACAAACCTCTGGACAAGTAGTAGCTGTTAAAAAATTTGTGGAATCTGAAGAT 120
QY 121 GATCCTGTGTTAAGAAAAATAGCACTAAGAGAAATACCTATGTTGAAGCAATTAACAAT 180
DB 121 GATCCTGTGTTAAGAAAAATAGCACTAAGAGAAATACCTATGTTGAAGCAATTAACAAT 180
QY 181 CCAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAAGGAAATTCGATTAGTTTTT 240
DB 181 CCAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAAGGAAATTCGATTAGTTTTT 240
QY 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGGAGTTGCTGAT 300
DB 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGGAGTTGCTGAT 300
QY 301 GGAGTGATCAAAAGCGTATTATGCAAAACACTTCAAGCTCTTAATTTCTGTCATATACAT 360
DB 301 GGAGTGATCAAAAGCGTATTATGCAAAACACTTCAAGCTCTTAATTTCTGTCATATACAT 360
QY 361 AACTGTATTACAGAGATATATAAACCTGAAATATTTCTTAATACTAAGCAAGGAATAATC 420
DB 361 AACTGTATTACAGAGATATATAAACCTGAAATATTTCTTAATACTAAGCAAGGAATAATC 420
QY 421 AAGATTGTGACTTCGGGTTTGCAAAATTTCTGATT 456
DB 421 AAGATTGTGACTTCGGGTTTGCAAAATTTCTGAGT 456

RESULT 13
AAD03815
ID AAD03815 standard; cDNA; 594 BP.

XX AAD03815;
AC AAD03815;
XX
XX
DT 19-JUN-2001 (first entry)
XX
DE Human kinase cDNA #4.
XX
XX Human; kinase; gene therapy; bioreactor; mental disorder;
KW biological disorder; ss.
KW

XX Homo sapiens.

XX
FH Key Location/Qualifiers
FI CDS 1..594
FI /*tag= a
FI /product= "Human kinase #4"
FI /note= "The coding region does not include stop codon"
FI /partial

XX WO200123579-A1.
XX
XX
PC 05-APR-2001.
XX
XX
PF 27-SEP-2000; 2000WO-US26621.
XX
XX
PR 28-SEP-1999; 99US-0156511.
XX
XX (LEXI-) LEXICON GENETICS INC.
XX
XX
XX Denoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
XX
XX WPI; 2001-266166/27.
DR P-PSDB; AAE00493.
XX
XX New isolated human kinase polynucleotide useful for generating
PT antibodies, as reagents in diagnostic assays and for screening for
PT compounds useful for treating mental, biological or medical diseases -
XX
XX Disclosure; Page 31; 38pp; English.
XX
XX The present sequence is a cDNA encoding novel human protein
CC (NHP) known as human kinase. The human kinases share structural
CC similarity with animal kinases, more particularly serine or
CC threonine protein kinases. Human kinase cDNA is useful for the
CC detection of mutant human kinase for the diagnosis of disease,
CC and also as a therapeutic. It is useful for screening drugs
CC effective in the treatment of symptomatic or phenotypic
CC manifestations perturbing the normal function of NHP in the
CC body. The NHP nucleotide sequences are useful for generation of
CC antibodies, as reagents in diagnostic assays, for the
CC identification of other cellular gene products related to human
CC kinases, and as reagents in assays for screening compounds that
CC are useful for treating mental, biological or medical disorders.
CC NHP oligonucleotides are used as probes. The labelled NHP probes
CC are useful for screening human genomic library for identifying
CC polymorphisms and as primers in amplification assays to detect
CC mutations within the exons, introns and splice sites that can
CC be used in diagnostics and pharmacogenomics. Nucleotide construct
CC encoding NHP products are used to genetically engineer cells
CC in vivo that functions as bioreactors in the body delivering a
CC continuous supply of NHP to the body. Nucleotide constructs
CC encoding functional NHPs are used in gene therapy for the
CC modulation of NHP expression.
XX
SQ Sequence 594 BP; 203 A; 94 C; 115 G; 182 T; 0 other;

Query Match 43.7%; Score 454.4; DB 22; Length 594;
Best Local Similarity 99.8%; Pred. No. 1.2e-111;
Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGAAAGTATGAAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGGTTGTATTCAA 60
DB 1 ATGGAAAGTATGAAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGGTTGTATTCAA 60
QY 61 TGCAGAAACAAACCTCTGGACAAGTAGTAGCTGTTAAAAAATTTGTGGAATCTGAAGAT 120
DB 61 TGCAGAAACAAACCTCTGGACAAGTAGTAGCTGTTAAAAAATTTGTGGAATCTGAAGAT 120
QY 121 GATCCTGTTGTTAAGAAAAATAGCACTAAGAGAAATACGTTATGTTGAAGCAATTAACAAT 180
DB 121 GATCCTGTTGTTAAGAAAAATAGCACTAAGAGAAATACGTTATGTTGAAGCAATTAACAAT 180
QY 181 CCAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAAGGAAATTCGATTAGTTTTT 240
DB 181 CCAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAAGGAAATTCGATTAGTTTTT 240
QY 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGGAGTTGCTGAT 300
DB 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGGAGTTGCTGAT 300
QY 301 GGAGTGATCAAAAGCGTATTATGCAAAACACTTCAAGCTCTTAATTTCTGTCATATACAT 360

Db 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTCTGTGATATACAT 360
QY 361 AACTGTATTACAGAGATATAAAACCTGAAATATTTCTAATACTAAGCAAGGAATATC 420
Db 361 AACTGTATTACAGAGATATAAAACCTGAAATATTTCTAATACTAAGCAAGGAATATC 420
QY 421 AAGATTGTGACTTCGGGTTTGACACAAATTTCTGATT 456
Db 421 AAGATTGTGACTTCGGGTTTGACACAAATTTCTGAGT 456

RESULT 14
ID AAF44670
XX AAF44670 standard; cDNA; 2615 BP.
AC AAF44670;
XX 27-MAR-2001 (first entry)
DT Novel protein kinase cDNA, SEQ ID NO: 50.
DE Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
XX immunosuppressive; cardiant; renal; antiinflammatory; antiasthmatic;
KW dermatological; antidiabetic; antiinfertility; gene therapy; vaccine;
KW immune disorder; cardiovascular disease; neurodegenerative disease;
KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;
KW inflammatory pelvic disease; multiple sclerosis; psoriasis; ss.
XX Mus musculus.

OS WO200073469-A2.
XX 07-DEC-2000.
XX 26-MAY-2000; 2000WO-US14842.
PF 28-MAY-1999; 99US-0136503.
PR (SUGEN-) SUGEN INC.
XX Plowman GD, Martinez R, Whyte D, Sudersanam S;
FA WPI; 2001-032161/04.
XX P-PSDB; AAB65643.

PT Nucleic acids encoding kinase polypeptides, useful for diagnosing and
PT treating immune-related diseases and disorders, cardiovascular disease,
PT neurodegenerative diseases and/or cancers -
XX Disclosure; Fig 2; 310pp; English.

PS The present sequence encodes a novel protein kinase. The nucleic acids
XX and the protein kinases they encode may be used in the treatment and
CC diagnosis of diseases associated with inappropriate kinase expression
CC such as immune-related diseases and disorders, cardiovascular disease,
CC neurodegenerative diseases and/or cancers. The nucleic acids and
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays. The kinase polypeptides may be used as antigens in the production
CC of antibodies of kinase expression and activity. Anti-kinase antibodies
CC and kinase antagonists may also be used to down regulate kinase
CC expression and activity. Diseases related to kinase expression and
CC activity include rheumatoid arthritis, atherosclerosis, autoimmune
CC disorders, complications of organ transplantation, myocardial infarction,
CC immune disorders, cardiomyopathies, strokes, renal failure,
CC oxidative-stress related disorders, chronic inflammatory bowel disease,
CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,
CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and
XX reproductive disorders.

SQ Sequence 2615 BP; 742 A; 580 C; 585 G; 708 T; 0 other;
Query Match 43.5%; Score 453.2; DB 22; Length 2615;

Best Local Similarity 73.1%; Pred. No. 4.6e-112;
Matches 658; Conservative 0; Mismatches 143; Indels 99; Gaps 2;
QY 135 GAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAACATCCAAATCTTGIGAA 194
Db 1 GAAAATAGCCCTCGGGAAATCCGTATGCTGAAG---TTGAAACACCCAAACCTCGTGAA 57
QY 195 CCTCATCGAGGTGTTTCAGGAGAAAAGGAAATGCAATTAGTTTGAATACTGTGATCA 254
Db 58 CCTCATCGAGGTGTTTCAGAAAGAAAGAGAAAGATGTCATCTAGTTTGTGACTGTGATCA 117
QY 255 TACACTTTTAATGAGCTGGAAGAAACCCAAATGGAGTTGCTGATGGAGTGATCAAAAG 314
Db 118 CACACTGTTAAACGAGCTGGAGAGAAACCCAAACGGAGTTTCTGATGGAGTGATTAAG 177
QY 315 CGTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGTCTATATACATAACTGTATTACAG 374
Db 178 TGTGCTATGGCAAAACCTTCAAGCCCTTAACCTTCTGTCAAGACACAATTTGATTCATCG 237
QY 375 AGATATATAAACCTGAAAATATTTCTAATACTAAGCAAGGAATAATCAAGATTTTGTGACTT 434
Db 238 GGATGTAAACCTGAAAACATCTCTAATAACCAAGCAAGGATGATAAAGATTTTGTGACTT 297
QY 435 CGGTTTGCACAAATTTCTGATTCAGGAGATGCCTACACCGATTATGTAGCTACAGATG 494
Db 298 TGAATTTGCACGAATTTCTAATTCAGGAGACGCCTACACAGACTATGTTGCCACAGGTG 357
QY 495 GTACCGAGCTCCTGAACTTCTTGTGGAGATATCTCAGTATGTTCTTTCAGTTCGATATATG 554
Db 358 GTACCGAGCCCCCGAACTTCTCGTGGGAGACACGAAGTACGGTTCTCTGTAGACGTGTG 417
QY 555 GGCTATTGTTGTGTTTTTGCAGAGCTCCTGACAGGCCAGCCACTGTGGCCTGGAAAATC 614
Db 418 GGCCGTGCGCTGTGTTTTTGACAGAGCTCCTGACGGGTACGCCACTCTGGCCGGAAAATC 477
QY 615 AGATGTGGACCAACTTTATCTGATAATCAGAACACTAGTAGAGACGGGTTTTCGCCATGT 674
Db 478 CGACGTGGACCAAGCTTTACCTGATCATCAGGACGTTG----- 514
QY 675 TGACCAGGCTGCTCGAACTCTTGACGTCAGTCAAGTGATCCACCTGCCGTAGCCTCTCAAAG 734
Db 515 ----- 514
QY 735 TGCTGGAATTACAGGAAATTAATCCCAAGACATCAATCAATCTTTTAAAGTAACGGGT 794
Db 515 -----GGGAAGCTGATTCCAAGACACCCAGTCTATCTTTAGGAGTAACCCAGTT 561
QY 795 TTTCATGGCATCAGTATACCTGAGCCAGAGACATGGAACACTCTTTGAGGAAAAGTTCTC 854
Db 562 TTTCCGCGGCATCAGCATACCTGACCAAGAGGACATGGAGACTCTTTGAAGAAAAATTTCTC 621
QY 855 AGATGTTTCATCTGTGGCTCTGAACTTTCATGAAGGGGTGCTGAAGATGAATCCAGATGA 914
Db 622 AAATGTTTCAGCTGTGGCTTTAAGTTTCATGAAGGGATGCCCTGAAGATGAATCTTGATGA 681
QY 915 CAGATTAACTGTTCCTCAACTCTGAGAGAGCTCCTACTTTGATTTCTTTTCAAGAGGCCCA 974
Db 682 GAGGCTGACCTGTGCCAGCTGTGGACAGTGCCTACTTTGAGTCTTTTCAAGAGGATCA 741
QY 975 AATTAAAAGAAAAGCAGCTAATGAAGGAAGAAAACAGAAAGCCCAACAGGTACTTCCGCT 1034
Db 742 AATGAAAAGAAAAGCCCGCAGTGAGGGGAGAAAGCCGAAGGCCCGCAGCAGATCAACTGCT 801

RESULT 15
ID ABZ77165
XX ABZ77165 standard; cDNA; 1086 BP.
AC ABZ77165;
XX 07-MAY-2003 (first entry)
XX Human protein kinase encoding cDNA SEQ ID NO:79.

XX Human; protein kinase; enzyme; antiasthmatic; antiinflammatory;
KW antidiabetic; antiparkinsonian; antimigraine; cardiant; cytostatic;
KW immunosuppressive; vulnerary; gene therapy; COPD; asthma; migraine;
KW chronic obstructive pulmonary disease; non-insulin dependent diabetes;
KW Parkinson's disease; myocardial infarction; inflammatory bowel disease;
KW autoimmune disorder; allograft rejection; graft versus host disease;
KW cancer; leukaemia; wound granulation; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1086
FT /*tag= a
FT /partial
FT /product= "protein kinase"
FT /note= "no start or stop codons given"
XX
PN WO200300901-A2.
XX
PN 03-JAN-2003.
XX
PF 24-JUN-2002; 2002WC-IR02358.
XX
PR 26-JUN-2001; 2001US-301098P.
PR 06-NOV-2001; 2001US-332870P.
XX
PA (DECO-) DECODE GENETICS EHF.
XX
XX Martinez RAM, Sigurdsson GT;
PI
XX
DR WPI; 2003-201429/19.
DR P-PSDB; ABP96087.
XX
XX New protein kinase genes and polypeptides, useful for diagnosing
PT diseases associated with a protein kinase, or in gene therapy for
PT treating e.g. Parkinson's disease, migraine, myocardial infarction,
PT allograft rejection or cancers -
XX
XX Claim 1; Page 89; 258pp; English.
PS
XX AB277126 to AB277165 encode the human protein kinases given in ABP96048
CC to ABP96087. The protein kinases have antiasthmatic, antiinflammatory,
CC antidiabetic, antiparkinsonian, antimigraine, cardiant, cytostatic,
CC immunosuppressive and vulnerary activities, and can be used in gene
CC therapy. A protein kinase therapeutic agent from the present invention,
CC particularly a protein kinase gene agonist or antagonist, can be used
CC for treating a disease or condition associated with a protein kinase in
CC an individual. These diseases include chronic obstructive pulmonary
CC diseases (COPD), asthma, non-insulin dependent diabetes, Parkinson's
CC disease, migraine, myocardial infarction, inflammatory bowel disease,
CC autoimmune disorders (e.g. allograft rejection or graft vs. host
CC disease), cancers (e.g. leukaemias) or wound granulation.
XX
SQ Sequence 1086 BP; 356 A; 227 C; 210 G; 293 T; 3 other;

Query Match 37.9%; Score 395; DB 25; Length 1086;
Best Local Similarity 83.9%; Pred. No. 1.2e-95;
Matches 501; Conservative 0; Mismatches 0; Indels 96; Gaps 1;

QY 292 GTTGCTGATCGAGTGATCAAAAGCGGTATTATGGCAAACACTCAAGCTCTTAATTCTGT 351
|||||
DB 283 GTTGCTGATGGAGTGATCAAAAGCGGTATTATGGCAAACACTCAAGCTCTTAATTCTGT 342
|||||

QY 352 CATATACATAACTGTATTACAGAGATATAAAACCTGAAATATTCTTAATCAAGCAA 411
|||||

DB 343 CATATACATAACTGTATTACAGAGATATAAAACCTGAAATATTCTTAATCAAGCAA 402
|||||

QY 412 GGAATAATCAAGATTGTGACTTCGGGTTTGCACAAATTCGATCCAGGAGATGCTTAC 471
|||||

DB 403 GGAATAATCAAGATTGTGACTTCGGGTTTGCACAAATTCGATCCAGGAGATGCTTAC 462
|||||

QY 472 ACCGATTATGTAGCTACGAGATGGTACCGAGCTCCTGAACTCTTTGTGGGAGATACTCAG 531
|||||

DB 463 ACCGATTATGTAGCTACGAGATGGTACCGAGCTCCTGAACTTCTTTGTGGGAGATACTCAG 522
|||||

QY 532 TATGGTCTTCAGTCGATATATGGGCTATTGGTTGTGTTTTCAGAGCTCCTGACAGGC 591
|||||

DB 523 TATGGTCTTCAGTCGATATATGGGCTATTGGTTGTGTTTTCAGAGCTCCTGACAGGC 582
|||||

QY 592 CAGCCACTGTGGCTGGAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTA 651
|||||

DB 583 CAGCCACTGTGGCTGGAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACT- 641
|||||

QY 652 GTAGAGACGGGTTTTCGCCATGTTGACCAGGCTGGTCTCGAACTCTTGACGTCGAAGTGAT 711
|||||

DB 642 ----- 641

QY 712 CCACCTGCCGTAGCTCTCAAAGTGTCTGGAATTACAGGAAATTAATCCCAAGACATCAA 771
|||||

DB 642 -----AGGAAATTAATCCCAAGACATCAA 666
|||||

QY 772 TCAATCTTTAAAGTAACGGGTTTTCATGGCATCAGTATACCTGAGCCAGAGACATG 831
|||||

DB 667 TCAATCTTTAAAGTAACGGGTTTTCATGGCATCAGTATACCTGAGCCAGAGACATG 726
|||||

QY 832 GAAACTCTTGAGGAAAGTTTCTCAGATGTTTCATCCTGTGGCTCTGAACCTTCATGAAG 888
|||||

DB 727 GAAACTCTTGAGGAAAGTTTCTCAGATGTTTCATCCTGTGGCTCTGAACCTTCATGAAG 783
|||||

Search completed: November 15, 2003, 16:54:28
Job time : 337.082 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2003, 11:42:01 ; Search time 4078.56 Seconds
(without alignments)
10441.655 Million cell updates/sec

Title: US-09-671-050-9
Perfect score: 1041
Sequence: 1 atggaagaagtatgaaaaatt.....aggtacttcggtcaaaagt 1041

Scoring table: IDENTITY_NUC
Gapop 10.0 ; Gapext 1.0

Searched: 2888711 seqs, 2045481336 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbi::*

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vi:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htg_mus:*
- 34: em_htg_pln:*
- 35: em_htg_rod:*
- 36: em_htg_mam:*
- 37: em_htg_vrt:*
- 38: em_sy:*
- 39: em_htgo_hum:*
- 40: em_htgo_mus:*
- 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1041	100.0	1041	6	AX107720	AX107720 Sequence
2	1025	98.5	1068	6	AX107714	AX107714 Sequence
3	839	80.6	945	6	AX107722	AX107722 Sequence
4	823	79.1	972	6	AX107716	AX107716 Sequence
5	762.8	73.3	1678	6	AX698820	AX698820 Sequence
6	757	72.7	1083	6	AX166534	AX166534 Sequence
7	709	68.1	1819	6	AX107724	AX107724 Sequence
8	645.6	62.0	687	6	AX286069	AX286069 Sequence
9	645.6	62.0	882	6	AX286067	AX286067 Sequence
10	537	51.6	911	6	AX056404	AX056404 Sequence
11	454.4	43.7	561	6	AX107712	AX107712 Sequence
12	454.4	43.7	594	6	AX107718	AX107718 Sequence
13	453.2	43.5	2615	6	AX056405	AX056405 Sequence
14	343.8	33.0	1363	9	HSSTHPKB	X66358 H.sapiens m
15	343.8	33.0	1612	6	AX399975	AX399975 Sequence
16	342.8	32.9	2952	3	AK112442	AK112442 Ciona int
17	251.8	24.2	1701	6	AR221995	AR221995 Sequence
18	251.8	24.2	3080	6	AR221994	AR221994 Sequence
19	251.8	24.2	3153	4	AB029045	AB029045 Oryctolag
20	243.8	23.4	1518	10	AB029067	AB029067 Mus muscu
21	243.8	23.4	3988	10	AB029066	AB029066 Mus muscu
22	243.8	23.4	4022	10	AB029065	AB029065 Mus muscu
23	241.2	23.2	1993	9	HSU35146	U35146 Homo sapien
24	202.6	19.5	1513	9	AF130372	AF130372 Homo sapi
25	202.6	19.5	1773	6	AX166533	AX166533 Sequence
26	202.6	19.5	2146	9	BC041799	BC041799 Homo sapi
27	201	19.3	1974	3	AF326965	AF326965 Trypanoso
28	200.6	19.3	1620	10	BC028871	BC028871 Mus muscu
29	199	19.1	80362	9	AC079615	AC079615 Homo sapi
30	191	18.3	1732	10	AF112183	AF112183 Rattus no
31	191	18.3	1738	10	AF112184	AF112184 Rattus no
32	186.6	17.9	101340	2	AC018104	AC018104 Drosophil
33	186.6	17.9	190668	3	AC007419	AC007419 Drosophil
34	186.6	17.9	260731	3	AE003611	AE003611 Drosophil
35	177.2	17.0	3399	9	AY217744	AY217744 Homo sapi
36	177.2	17.0	3399	9	HSTXP3	Y15057 Homo sapien
37	177.2	17.0	3441	9	BC036091	BC036091 Homo sapi
38	175.2	16.8	1866	6	AX056407	AX056407 Sequence
39	166	15.9	960	3	TCU69958	U69958 Trypanosoma
40	163.8	15.7	903	8	AF026545	AF026545 Pneumocys
41	162	15.6	1476	6	AR087351	AR087351 Sequence
42	162	15.6	1476	6	AR281916	AR281916 Sequence
43	162	15.6	1476	9	HSCDK2MR	X61622 H.sapiens C
44	162	15.6	1635	6	A84497	A84497 Sequence 11
45	162	15.6	1635	6	A84499	A84499 Sequence 11

ALIGNMENTS

RESULT 1	AX107720	AX107720	1041 bp	DNA	Linear	PAT 30-APR-2001
LOCUS	Sequence 9 from Patent WO0123579.					
DEFINITION	AX107720					
ACCESSION	AX107720.1	GI:13923201				
VERSION						
KEYWORDS						
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
TITLE	Doncho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and Sands,A.T.					
	Human kinase proteins and polynucleotides encoding the same					

JOURNAL Patent: WO 0123579-A 9 05-APR-2001;
Lexicon Genetics Incorporated (US)
FEATURES Location/Qualifiers
Source 1..104;

BASE COUNT	344 a	190 c	228 g	279 t
ORIGIN				

```
Query Match      100.0%; Score 1041; DB 6; Length 1041;
Best Local Similarity 100.0%; Pred. NO. 2.9e-242;
Matches 1041; Conservative 0; Mismatches 0; Indels 0;
```

QY	1	ATGGA	AAAGTA	TGAAAA	AATTAGCT	AAAGACTGGAGA	AGGGTCTTAT	TGGGGTTG	TATTCAAA	60		
DB	1	ATGGA	AAAGTA	TGAAAA	AATTAGCT	AAAGACTGGAGA	AGGGTCTTAT	TGGGGTTG	TATTCAAA	60		
QY	61	TGCAG	AAACAA	AAACCTCT	TGGACA	AGTAGT	AGCTGT	TAAAAA	ATTTGTG	GAATCTCAAGAT	120	
DB	61	TGCAG	AAACAA	AAACCTCT	TGGACA	AGTAGT	AGCTGT	TAAAAA	ATTTGTG	GAATCTCAAGAT	120	
QY	121	GATCC	TGTTGT	TAAGAA	ATAGCA	CTAAGAG	AAATAC	GTATGTTGA	GCAATTTAAAA	ACAT	180	
DB	121	GATCC	TGTTGT	TAAGAA	ATAGCA	CTAAGAG	AAATAC	GTATGTTGA	GCAATTTAAAA	ACAT	180	
QY	181	CCAAAT	CTTGTG	AAACCTCAT	CGAGG	TGTTTC	AGGAG	AAATTC	AGGAAAT	TSCATTTAGTTT	240	
DB	181	CCAAAT	CTTGTG	AAACCTCAT	CGAGG	TGTTTC	AGGAG	AAATTC	AGGAAAT	TSCATTTAGTTT	240	
QY	241	GAATAC	TGTGAT	CATACAC	TCTTTT	AAATGAG	CTGGAA	AGAAACCC	AAATGGAG	TTCCTGAT	300	
DB	241	GAATAC	TGTGAT	CATACAC	TCTTTT	AAATGAG	CTGGAA	AGAAACCC	AAATGGAG	TTCCTGAT	300	
QY	301	GGAGT	GATCA	AAAGCGT	ATTAT	TGGCA	AAACACTT	CAAGCTCT	TAAATTTCTG	TCAATACAT	360	
DB	301	GGAGT	GATCA	AAAGCGT	ATTAT	TGGCA	AAACACTT	CAAGCTCT	TAAATTTCTG	TCAATACAT	360	
QY	361	AAC	TGATTCA	CAGAGAT	ATAAAA	ACCTG	AAAAATAT	TCTATA	AACTAAGCA	AGGAATATC	420	
DB	361	AAC	TGATTCA	CAGAGAT	ATAAAA	ACCTG	AAAAATAT	TCTATA	AACTAAGCA	AGGAATATC	420	
QY	421	AAGATT	TGACTT	CGGGTT	TGCAC	AAATTC	TGATTC	CCAGG	AGATGCCT	ACACCGATTAT	480	
DB	421	AAGATT	TGACTT	CGGGTT	TGCAC	AAATTC	TGATTC	CCAGG	AGATGCCT	ACACCGATTAT	480	
QY	481	GTAGCT	ACGAGAT	GGTAC	CGAGCT	CCTG	AACTTCTT	TGTGG	GAGATACT	CAGTATGGTTCT	540	
DB	481	GTAGCT	ACGAGAT	GGTAC	CGAGCT	CCTG	AACTTCTT	TGTGG	GAGATACT	CAGTATGGTTCT	540	
QY	541	TCAGT	CGATAT	TATGG	CTATT	TGGTT	TGTGTTT	TGCG	AGACTCT	TGACAGGCCACCTG	600	
DB	541	TCAGT	CGATAT	TATGG	CTATT	TGGTT	TGTGTTT	TGCG	AGACTCT	TGACAGGCCACCTG	600	
QY	601	TGGCCT	TGGAAA	ATCAG	ATGTG	GACCA	ACTTTAT	CTGTAT	ATCAGAA	CACTAGTAGAGACG	660	
DB	601	TGGCCT	TGGAAA	ATCAG	ATGTG	GACCA	ACTTTAT	CTGTAT	ATCAGAA	CACTAGTAGAGACG	660	
QY	661	GGGTT	TCGCC	ATGTTG	ACCAG	GGCTGGT	CTCGAA	ACTCTT	TGACGT	CAAGTGATCCACCTGCC	720	
DB	661	GGGTT	TCGCC	ATGTTG	ACCAG	GGCTGGT	CTCGAA	ACTCTT	TGACGT	CAAGTGATCCACCTGCC	720	
QY	721	GTAGC	CTCTCA	AAAGT	GCTGGA	ATTAC	AGGAAA	ATTAAT	CCCAAG	ACATCAATCTTT	780	
DB	721	GTAGC	CTCTCA	AAAGT	GCTGGA	ATTAC	AGGAAA	ATTAAT	CCCAAG	ACATCAATCTTT	780	
QY	781	AAAAG	TAA	CGGGTTT	TTCAT	TGGCAT	TCAGT	ATACCT	TGAGCC	CAGAGACATGGAA	840	
DB	781	AAAAG	TAA	CGGGTTT	TTCAT	TGGCAT	TCAGT	ATACCT	TGAGCC	CAGAGACATGGAA	840	
QY	841	GAGG	AAAAG	TTCTC	AGATGTT	TCAT	CTCTG	CGCTCT	TGAAC	CTTTC	TGAAGGGGTGCTGAAG	900
DB	841	GAGG	AAAAG	TTCTC	AGATGTT	TCAT	CTCTG	CGCTCT	TGAAC	CTTTC	TGAAGGGGTGCTGAAG	900

QY	901	ATGAATCCAGATGACAGATTAAACCTGTTCCCAACTCCTGGAGAGCTCCTACTTTGATTCT	960
DB	901	ATGAATCCAGATGACAGATTAAACCTGTTCCCAACTCCTGGAGAGCTCCTACTTTGATTCT	960
QY	961	TTTCAAGAGGCCCCCAATTAAAAGAAAGCACGTAATGAAGGAAGAAACAGAAGACGCCAA	1020
DB	961	TTTCAAGAGGCCCCCAATTAAAAGAAAGCACGTAATGAAGGAAGAAACAGAAGACGCCAA	1020
QY	1021	CAGGTACTTCCGGTCAAAAAGT	1041
DB	1021	CAGGTACTTCCGGTCAAAAAGT	1041

RESULT 2
AX107714
LOCUS AX107714 1068 bp
DEFINITION Sequence 3 from Patent WO0123579.

REFERENCE	Donoho, G., Turner, C.A., Nehls, M., Friedrich, G., Zambrowicz, B. and Sands, A.T.
AUTHORS	Human kinase proteins and polynucleotides encoding the same
TITLE	Patent: WO 0123579-A 3 05-APR-2001;
COUNCIL	Lexicon Genetics Incorporated (US)

BASE COUNT	344 a	197 c	240 g	287 t
ORIGIN				

Query Match	98.5%	Score 1025;	DB 6;	Length 1068;
Best Local Similarity	100.0%	Pred. No. 2.2e-238;		
Matches:1025;	Conservative	C;	Mismatches	0;
			Indels	0

Cy	1	ATGGAAGAAGTATGAAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGGTTGTATTCAAA	60
Dh	1	ATGGAAGAAGTATGAAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGGTTGTATTCAAA	60
Cy	61	TGCAGAAACAAAACCTCTGGACAAGTAGTAGCTGTAAAAAAATTTGTGGAATCTGAAGAT	120
Dh	61	TGCAGAAACAAAACCTCTGGACAAGTAGTAGCTGTAAAAAAATTTGTGGAATCTGAAGAT	120
Cy	121	GATCCTGTGTTAAGAAAAATAGCACTAAGAGAAATACGTATGTTGAAGCCAATTAAAAACAT	180
Dh	121	GATCCTGTGTTAAGAAAAATAGCACTAAGAGAAATACGTATGTTGAAGCCAATTAAAAACAT	180
Cy	191	CCAAATCTGTGAACCTCATTCGAGGTGTTCCGGAGAAAAAGGAAAAATGCATTTAGTTTTT	240
Dh	191	CCAAATCTGTGAACCTCATTCGAGGTGTTCCGGAGAAAAAGGAAAAATGCATTTAGTTTTT	240
Cy	241	GAATACTGTGATCATACACTTTTAAATGAGCTGGAAGAAAAACCCAAATGGAGTTGCTGAT	300
Dh	241	GAATACTGTGATCATACACTTTTAAATGAGCTGGAAGAAAAACCCAAATGGAGTTGCTGAT	300
Cy	301	GGAGTGATCAAAAGCGTATTATGGCAACACCTTCAAGCTCTTAATTTCTGTGATATACAT	360
Dh	301	GGAGTGATCAAAAGCGTATTATGGCAACACCTTCAAGCTCTTAATTTCTGTGATATACAT	360
Cy	361	AACCTGATTCAAGAGATATAAAACCTGAAAAATTTCTTAATAACTAAGCAAGGAATAATC	420
Dh	361	AACCTGATTCAAGAGATATAAAACCTGAAAAATTTCTTAATAACTAAGCAAGGAATAATC	420
Cy	421	AAGATTTGTGACTTCGGGTTTTGCCAAAAATTTCTGATTCAGGAGATGCCTACACCGATTAT	480

Db 421 AAGATTGTGACTTCGGGTTTGACAAATCTGTATCCAGGAGATGCCTACACCGATTAT 480
QY 481 GTAGCTACGAGATGGTACCGAGCTCCTGAACTTCTTGTGGGAGATACCTCAGTATGTTCT 540
Db 481 GTAGCTACGAGATGGTACCGAGCTCCTGAACTTCTTGTGGGAGATACCTCAGTATGTTCT 540
QY 541 TCAGTCGATATATGGGCTATTGGTTGGTCTTTTSCAGAGCTCCTGACAGGCCAGCCACTG 600
Db 541 TCAGTCGATATATGGGCTATTGGTTGGTCTTTTSCAGAGCTCCTGACAGGCCAGCCACTG 600
QY 601 TGGCCTGGAAATCAGATGTGGACCACTTTATCTGATATCAGAAACACTAGTAGAGACG 660
Db 601 TGGCCTGGAAATCAGATGTGGACCACTTTATCTGATATCAGAAACACTAGTAGAGACG 660
QY 661 GGGTTTCGCCATGTTGACCAAGGCTGGTCTCGAACTCTTGACGTCAAGTGATCCACTGCC 720
Db 661 GGGTTTCGCCATGTTGACCAAGGCTGGTCTCGAACTCTTGACGTCAAGTGATCCACTGCC 720
QY 721 GTAGCCTCTCAAGTGTCTGGAATTACAGGAAATTAATCCCAAGACATCAATCAATCTTT 780
Db 721 GTAGCCTCTCAAGTGTCTGGAATTACAGGAAATTAATCCCAAGACATCAATCAATCTTT 780
QY 781 AAAGTAAACGGGTTTTCATGGCATGGCATACCTGAGCCAGACACATGGAAACTCTT 840
Db 781 AAAGTAAACGGGTTTTCATGGCATGGCATACCTGAGCCAGACACATGGAAACTCTT 840
QY 841 GAGGAAAGTTCTCAGATGTTTCCATGGCATGGCATACCTGAGCCAGACACATGGAAACTCTT 900
Db 841 GAGGAAAGTTCTCAGATGTTTCCATGGCATGGCATACCTGAGCCAGACACATGGAAACTCTT 900
QY 901 ATGAATCCAGATGACAGATGTTTCCATGGCATGGCATACCTGAGCCAGACACATGGAAACTCTT 960
Db 901 ATGAATCCAGATGACAGATGTTTCCATGGCATGGCATACCTGAGCCAGACACATGGAAACTCTT 960
QY 961 TTTCAAGAGGCCCAAAATTAAGAAAGCACGTAATGAAGGAAAGAAACAGAGACGCCAA 1020
Db 961 TTTCAAGAGGCCCAAAATTAAGAAAGCACGTAATGAAGGAAAGAAACAGAGACGCCAA 1020
QY 1021 CAGGT 1025
Db 1021 CAGGT 1025

RESULT 3
AX107722
LOCUS AX107722 945 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 11 from Patent WO0123579.
ACCESSION AX107722
VERSION AX107722.1 GI:13923202
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and Sands,A.T.
TITLE Human kinase proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0123579-A 11 05-APR-2001;
Lexicon Genetics Incorporated (US)
FEATURES
source 1..945
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 323 a 165 c 202 g 255 t
ORIGIN
Query Match 80.6%; Score 839; DB 6; Length 945;
Best Local Similarity 90.8%; Pred. No. 3.3e-193;
Matches 945; Conservative 0; Mismatches 0; Indels 96; Gaps 1;
QY 1 ATGGAAGATATGAAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGGTTGTATTCAA 60

Db 1 ATGGAAGATATGAAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGGTTGTATTCAA 60
QY 61 TGCAGAAACAAACCTCTGACAAAGTAGTAGCTTTAAAAAATTTGTGGAATCTGAAGAT 120
Db 61 TGCAGAAACAAACCTCTGACAAAGTAGTAGCTTTAAAAAATTTGTGGAATCTGAAGAT 120
QY 121 GATCCTGTTGTTAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAACAT 180
Db 121 GATCCTGTTGTTAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAACAT 180
QY 181 CCAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAAGGAAATGCATTTAGTTT 240
Db 181 CCAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAAGGAAATGCATTTAGTTT 240
QY 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATCGAGTTGCTGAT 300
Db 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATCGAGTTGCTGAT 300
QY 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGCATATACAT 360
Db 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGCATATACAT 360
QY 361 AACTGTATTACAGAGATATAAAACCTGAAATATTTCTAATACTAAGCAAGGAATAATC 420
Db 361 AACTGTATTACAGAGATATAAAACCTGAAATATTTCTAATACTAAGCAAGGAATAATC 420
QY 421 AAGATTTGTGACTTCGGGTTTGCAACAAATTTCTGATTCAGGAGATGCCCTACACCGATTAT 480
Db 421 AAGATTTGTGACTTCGGGTTTGCAACAAATTTCTGATTCAGGAGATGCCCTACACCGATTAT 480
QY 481 GTAGTACGAGATGGTACCGAGCTCCTGAACTTCTTGTGGAGATACCTCAGTATGTTCT 540
Db 481 GTAGTACGAGATGGTACCGAGCTCCTGAACTTCTTGTGGAGATACCTCAGTATGTTCT 540
QY 541 TCAGTCGATATATGGGCTATTGGTGTGTTTTCAGAGCTCCTGACAGCCAGCCACTG 600
Db 541 TCAGTCGATATATGGGCTATTGGTGTGTTTTCAGAGCTCCTGACAGCCAGCCACTG 600
QY 601 TGGCCTGGAAATCAGATGTGGACCAACTTTATCTGATAATCAGAAACACTAGTAGAGACG 660
Db 601 TGGCCTGGAAATCAGATGTGGACCAACTTTATCTGATAATCAGAAACACTAGTAGAGACG 660
QY 661 GGGTTTCGCCATGTTGACCAAGGCTGGTCTCGAACTCTTGACGTCAAGTGATCCACTGCC 720
Db 651 -----
QY 721 GTAGCCTCTCAAAGTGTCTGGAATTACAGGAAATTAATCCCAAGACATCAATCAATCTTT 780
Db 651 -----AGGAAATTAATCCCAAGACATCAATCAATCTTT 684
QY 781 AAAAGTAAACGGGTTTTCATGGCATGGCATACCTGAGCCAGACACATGGAAACTCTT 840
Db 685 AAAAGTAAACGGGTTTTCATGGCATGGCATACCTGAGCCAGACACATGGAAACTCTT 744
QY 841 GAGGAAAGTTCTCAGATGTTTCCATGGCTCTGAACTTTCATGAGGGGTGCTCTGAAG 900
Db 745 GAGGAAAGTTCTCAGATGTTTCCATGGCTCTGAACTTTCATGAGGGGTGCTCTGAAG 804
QY 901 ATGAATCCAGATGACAGATTAACCTGTTCCCAACTCCTGGAGAGCTCCTACTTTGATTCT 960
Db 805 ATGAATCCAGATGACAGATTAACCTGTTCCCAACTCCTGGAGAGCTCCTACTTTGATTCT 864
QY 961 TTTCAAGAGGCCCAAAATTAAGAAAGCACGTAATGAAGGAAAGAAACAGAGACGCCAA 1020
Db 865 TTTCAAGAGGCCCAAAATTAAGAAAGCACGTAATGAAGGAAAGAAACAGAGACGCCAA 924
QY 1021 CAGGTACTTCCGCTCAAAGT 1041
Db 925 CAGGTACTTCCGCTCAAAGT 945

AX107716
LOCUS AX107716 972 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 5 from Patent WO0123579.
ACCESSION AX107716
VERSION AX107716.1 GI:13923199
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and Sands,A.
TITLE Human kinase proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0123579-A 5 05-APR-2001;
Lexicon Genetics Incorporated (US);
FEATURES
source Location/Qualifiers
1. 972
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 323 a 172 c 214 g 263 t
ORIGIN

Query Match 79.1%; Score 823; DB 6; Length 972;
Best Local Similarity 90.6%; Pred. No. 2.5e-189;
Matches 929; Conservative 0; Mismatches 0; Indels 96; Gaps 1;

QY 1 ATGGAAAGTATGAAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGGTGTATTCAA 60
DB 1 ATGGAAAGTATGAAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGGTGTATTCAA 60

QY 61 TGCAGAAACAAACCTCTGGACAGTAGTAGCTGTTAAAAAATTGTGGAACTCTGAAGAT 120
DB 61 TGCAGAAACAAACCTCTGGACAGTAGTAGCTGTTAAAAAATTGTGGAACTCTGAAGAT 120

QY 121 GATCCTGTGTTAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAAACAT 180
DB 121 GATCCTGTGTTAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAAACAT 180

QY 181 CCAATCTGTGTAACCTCATCGAGGTGTTTCAGGAGAAAAAGGAAATGCATTAGTTTT 240
DB 181 CCAATCTGTGTAACCTCATCGAGGTGTTTCAGGAGAAAAAGGAAATGCATTAGTTTT 240

QY 241 GAATCTGTGATCATACCTTTTAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT 300
DB 241 GAATCTGTGATCATACCTTTTAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT 300

QY 301 GGAGTGATCAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAAATTTCTGTCAATACAT 360
DB 301 GGAGTGATCAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAAATTTCTGTCAATACAT 360

QY 361 AACTGTATTACAGAGATATAAAACCTGAAATATTCTAATAACTAAGCAAGGAATAATC 420
DB 361 AACTGTATTACAGAGATATAAAACCTGAAATATTCTAATAACTAAGCAAGGAATAATC 420

QY 421 AAGATTGTGACTTCGGGTTTGCAAAAATTCTGATTCAGGAGATGCCTACACCGATTAT 480
DB 421 AAGATTGTGACTTCGGGTTTGCAAAAATTCTGATTCAGGAGATGCCTACACCGATTAT 480

QY 481 GTAGCTACGAGATGGTACCGAGCTCCTGAACTTCTGTGGAGATACCTCAGTATGGTCT 540
DB 481 GTAGCTACGAGATGGTACCGAGCTCCTGAACTTCTGTGGAGATACCTCAGTATGGTCT 540

QY 541 TCAGTCGATATATGGGCTATTGGTTGTGTTTTCAGAGCTCCTGACAGGCCAGCCACTG 600
DB 541 TCAGTCGATATATGGGCTATTGGTTGTGTTTTCAGAGCTCCTGACAGGCCAGCCACTG 600

QY 601 TGGCCTGGAAATACAGATGTGGACCACTTTATCTGATAATCAGAACACTAGTAGAGAG 660
DB 601 TGGCCTGGAAATACAGATGTGGACCACTTTATCTGATAATCAGAACACTAGTAGAGAG 660

QY 661 GGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTCTTGACGTCAGTGCACCTGCC 720

DB 651 ----- 650
QY 721 GTAGCCTCTCAAGTGTCTGGAATTACAGGAAAAATTAAATCCCAAGACATCAATCAATCTTT 780
DB 651 -----AGGAAAAATTAAATCCCAAGACATCAATCAATCTTT 684

QY 781 AAAAGTAAACGGGTTTTTCCATGGCATCAGTATACCTGAGCCAGAGACATGGAAACTCTT 840
DB 685 AAAAGTAAACGGGTTTTTCCATGGCATCAGTATACCTGAGCCAGAGACATGGAAACTCTT 744

QY 841 GAGGAAAAAGTTCTCAGATGTTTCATCTCTGTGGCTCTGAACTTCATGAAGGGGTGCTGAAG 900
DB 745 GAGGAAAAAGTTCTCAGATGTTTCATCTCTGTGGCTCTGAACTTCATGAAGGGGTGCTGAAG 804

QY 901 ATGAATCCAGATGACAGATTAACTGTTCCCAACTCCTGGAGAGCTCCTACTTTGATTCT 960
DB 805 ATGAATCCAGATGACAGATTAACTGTTCCCAACTCCTGGAGAGCTCCTACTTTGATTCT 864

QY 961 TTTCAAGAGGCCCAAAATTAAAGAAAGCACCGTAATGAAGGAAGAAACAGAGACGCCAA 1020
DB 865 TTTCAAGAGGCCCAAAATTAAAGAAAGAACACGTAATGAAGGAAGAAACAGAGACGCCAA 924

QY 1021 CAGGT 1025
DB 925 CAGGT 929

RESULT 5
AX698820
LOCUS AX698820 1678 bp DNA linear PAT 02-APR-2003
DEFINITION Sequence 6 from Patent WO02079473.
ACCESSION AX698820
VERSION AX698820.1 GI:29499608
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Panzer,S.R., Lincoln,S.E., Altus,C.M., Dufour,G.E., Hillman,J.L., Jones,A.L., Dam,T.C., Liu,T.F., Harris,B., Flores,V., Daffo,A., Marwaha,R., Chen,A.J., Chang,S.C., Gerstin,J.E., Peralta,C.H., David,N.H. and Lewis,S.A.
TITLE Molecules for diagnostics and therapeutics
JOURNAL: Patent: WO 02079473-A 6 10-OCT-2002;
Incyte Genomics, Inc. (US)
FEATURES
Location/Qualifiers
1. 1678
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: LI:058298.1:2001JAN12"
BASE COUNT 541 a 308 c 336 g 493 t
ORIGIN

Query Match 73.3%; Score 762.8; DB 6; Length 1678;
Best Local Similarity 90.0%; Pred. No. 1e-174;
Matches 925; Conservative 0; Mismatches 2; Indels 101; Gaps 6;

QY 1 ATGGAAAAGTATGAAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGGTGTATTCAAA 60
DB 131 ATGGAGAAGTATGAAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGGTGTATTCAAA 190

QY 61 TGCAGA-AACAAAACCTCTGGACAAGTAGTAGCTGTTAAAAAATTGTGGAATCTGAAGA 119
DB 191 TGCAGATAACAAAACCTCTGGACAAGTAGTAGCTGTTAAAAAATTGTGGAATCTGAAGA 250

QY 120 TGATCCTGTTGTTAAGAAAAATAGCACTAAGAGAAATACGTA-TGTTGAAGCAATTAAAC 178
DB 251 TGATCCTGTTGTTAAGAAAAATAGCACTAAGAGAAATACGTA-TGTTGAAGCAATTAAAC 310

QY 179 ATCCAAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAAGGAAATGCATTAGTTT 238

|||||
Db 311 ATCCAAATCTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAAGGAAATGCAATTAGTTT 370
Qy 239 TTGAATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTG 298
Db 371 TTGAATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTG 430
Qy 299 ATGGAGTGATCAANAGCGTATTATGGCAACACCTTCAAGCTCTTAAATTTCTGCATATAC 358
Db 431 ATGGAGTGATCAANAGCGTATTATGGCAACACCTTCAAGCTCTTAAATTTCTGCATATAC 490
Qy 359 ATAACTGTATTACAGAGATATAAAGACCTGAAATATTCTAATACTAAGCAAGGAATAA 418
Db 491 ATAACTGTATTACAGAGATATAAAGACCTGAAATATTCTAATACTAAGCAAGGAATAA 550
Qy 419 TCAAGATTGTGACTTCGGGTTTGCACAAATCTGATTCAGGAGATGCCATACACCGATT 478
Db 551 TCAAGATTGTGACTTCGGGTTTGCACAAATCTGATTCAGGAGATGCCATACACCGATT 610
Qy 479 ATGTAGCTA-CGAGATGGTACCGAGCT-CCTGAACTTCTGTTGGGAGATCT-CAGTATG 535
Db 611 ATGTAGCTAGCGAGATGGTACCGAGCTCCCTGAACTTCTGTTGGGAGATCTCCAGTATG 670
Qy 536 GTTCTTCAGTCGATATATGGGCTATTGTTGTTGTTTTCAGAGCTCCTGCACAGGCCAGC 595
Db 671 GTTCTTCAGTCGATATATGGGCTATTGTTGTTGTTTTCAGAGCTCCTGCACAGGCCAGC 730
Qy 596 CACTGTGGCCTGGAAATCAGATGTGGACCAACTTTATCTGTATATCAGACACTAGTAG 555
Db 731 CACTGTGGCCTGGAAATCAGATGTGGACCAACTTTATCTGTATATCAGACACT----- 785
Qy 656 AGACGGGTTTCGCCATGTTCAGCAGGCTGTCTCGAACTCTTGACGTCAAGTGATCCAC 715
Db 786 ----- 785
Qy 716 CTGCCGTAGCCTCTCAAAGTGCTGGAATTACAGGAAATTAATCCCAAGACATCAATCAA 775
Db 786 -----AGGAAATTAATCCCAAGACATCAATCAA 814
Qy 776 TCTTTAAAGTAACGGGTTTTTCCATGGCATCAGTATACCTGAGCCAGACATGGAAA 835
Db 815 TCTTTAAAGTAACGGGTTTTTCCATGGCATCAGTATACCTGAGCCAGACATGGAAA 874
Qy 836 CTCTTCAGGAAAGTTCTCAGATGTTTCATCTGTGGCTCTGAACTTCATGAAGGGGTGTC 895
Db 875 CTCTTCAGGAAAGTTCTCAGATGTTTCATCTGTGGCTCTGAACTTCATGAAGGGGTGTC 934
Qy 896 TGAAGATGAATCCAGATGACAGATTAACTGTTCCCACTCCTGGAGAGCTCCTACTTTG 955
Db 935 TGAAGATGAATCCAGATGACAGATTAACTGTTCCCACTCCTGGAGAGCTCCTACTTTG 994
Qy 956 ATTCTTTTCAAGAGGCCCAATTAAGAAAGAAAGCAACGTAATGAAGGAAGAAACAGAAGAC 1015
Db 995 ATTCTTTTCAAGAGGCCCAATTAAGAAAGAAAGCAACGTAATGAAGGAAGAAACAGAAGAC 1054
Qy 1016 GCCAACAG 1023
Db 1055 GCCAACAG 1062

RESULT 6
AX166534
LOCUS
DEFINITION
Sequence 25 from Patent WO0138503.
ACCESSION
AX166534
VERSION
AX166534.1 GI:14546879
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS
Plowman,G.D., Whyte,D., Manning,G.S., Sudarsanam,S.S., Martinez,R.,

Flanagan,P. and Clary,D.S.
Novel human protein kinases and protein kinase-like enzymes
Patent: WO 0138503-A 25 31-MAY-2001;
Sugen, Inc. (US)
Location/Qualifiers
1..1083
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
366 a 199 c 225 g 293 t
BASE COUNT
ORIGIN
Query Match 72.7%; Score 757; DB 6; Length 1083;
Best Local Similarity 86.1%; Pred. No. 2.7e-173;
Matches 927; Conservative 0; Mismatches 0; Indels 150; Gaps 2;
Qy 1 ATGGAAAAGTATGAAAAATTTAGCTAAGACTGGAGAAGGGTCTTATGGGGTTGTATTCAA 60
Db 1 ATGGAAAAGTATGAAAAATTTAGCTAAGACTGGAGAAGGGTCTTATGGGGTTGTATTCAA 60
Qy 61 TGCAGAAACAAACCTCTCGACAAAGTAGTAGCTGTAAAAAAATTTGTGGAATCTGAAGAT 120
Db 61 TGCAGAAACAAACCTCTCGACAAAGTAGTAGCTGTAAAAAAATTTGTGGAATCTGAAGAT 120
Qy 121 GATCCTGTTGTTAAGAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTTAAACAT 180
Db 121 GATCCTGTTGTTAAGAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTTAAACAT 180
Qy 181 CCAAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAAGGAAAAATGCAATTTAGTTT 240
Db 181 CCAAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAAGGAAAAATGCAATTTAGTTT 240
Qy 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGGAGTTGCTGAT 300
Db 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGGAGTTGCTGAT 300
Qy 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAAATTTCTGCATATACAT 360
Db 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAAATTTCTGCATATACAT 360
Qy 361 AACTGTATTCAAGAGATATAAAACCTGAAAAATATTCTAATAACTAAGCAAGGAATAATC 420
Db 361 AACTGTATTCAAGAGATATAAAACCTGAAAAATATTCTAATAACTAAGCAAGGAATAATC 420
Qy 421 AAGATTGTGACTTCGGGTTTGCACAAATTTCTGA----- 454
Db 421 AAGATTGTGACTTCGGGTTTGCACAAATTTCTGAATAATTCACATGTTGGGAGAAATTGAC 480
Qy 455 -----TTCCAGGAGATGCCTACACCGATTATGTAGCT 486
Db 481 TTACCTGATCTGATTGATGCATTTGCAAGTTCCAGGAGATGCCTACACCGATTATGTAGCT 540
Qy 487 ACGAGATGGTACCGAGCTCCTGAACCTTCTTGTGGAGATACTCAGTATGGTTCTTCAGTC 546
Db 541 ACGAGATGGTACCGAGCTCCTGAACCTTCTTGTGGAGATACTCAGTATGGTTCTTCAGTC 600
Qy 547 GATATATGGGCTATTGTTGTTTTCAGAGCTCCTGACAGCCAGCCACTGTGGCCT 606
Db 601 GATATATGGGCTATTGTTGTTTTCAGAGCTCCTGACAGCCAGCCACTGTGGCCT 660
Qy 607 GGAAAAATCAGATGTGGACCAACTTTTATCTGATAATCAGAAACACTAGTAGACACGGGTTT 666
Db 661 GGAAAAATCAGATGTGGACCAACTTTTATCTGATAATCAGAAACACT----- 704
Qy 667 CGCCATGTTGACCCAGGCTGGTCTCGAACTTTGACGTCAAGTGATCCACCTGCCGTAGCC 726
Db 705 ----- 704
Qy 727 TCTCAAAGTGTGGAATTACAGGAAATTAATCCCAAGACATCAATCAATCTTTTAAAGT 786
Db 705 -----AGGAAATTAATCCCAAGACATCAATCAATCTTTTAAAGT 744
Qy 787 AACGGGTTTTTCCATGGCATCAGTATACCTGAGCCAGAAAGACATGGAAACTCTTGAGGAA 846

Db 745 AACGGGTTTTCCATGGCATCAGTATACCTGAGCCAGAGACATGGAACCTCTTGAGGAA 804
QY 847 AAGTTCTCAGATGTTTCATCCTGTGGCTCTGAACTTCATGAAGGGGTGTCTGAAGATGAAT 906
Db 805 AAGTTCTCAGATGTTTCATCCTGTGGCTCTGAACTTCATGAAGGGGTGTCTGAAGATGAAT 864
QY 907 CCAGATGACAGATTACCTGTTCCTCAACTCCTGGAGAGCTCCTACTTTGATTTCTTTTCAA 966
Db 865 CCAGATGACAGATTACCTGTTCCTCAACTCCTGGAGAGCTCCTACTTTGATTTCTTTTCAA 924
QY 967 GAGGCCCCAAATTAAAGAAAGACCGTAAATGAAGGAAGAAACAGAGAGCGCCACACG 1023
Db 925 GAGGCCCCAAATTAAAGAAAGACCGTAAATGAAGGAAGAAACAGAGAGCGCCACACG 981

RESULT 7
AX107724
LOCUS AX107724 1819 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 13 from Patent WO0123579.
ACCESSION AX107724
VERSION AX107724.1 GI:13923203
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1
AUTHORS Denoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and Sands,A.T.
TITLE Human kinase proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0123579-A 13 05-APR-2001.
FEATURES
source Location/Qualifiers
1..1819
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 539 a 368 c 406 g 505 t 1 others
ORIGIN

Query Match 68.1%; Score 709; DB 6; Length 1819;
Best Local Similarity 77.0%; Pred. No. 1.2e-161;
Matches 1025; Conservative 0; Mismatches 0; Indels 306; Gaps 1;

QY 1 ATGGAAGAAGTATGAAAAATTAGCTAAGACTGGAGAAGGGTCTTATGGGTGTGATTTCAA 60
Db 138 ATGGAAGAAGTATGAAAAATTAGCTAAGACTGGAGAAGGGTCTTATGGGTGTGATTTCAA 197
QY 61 TGCAGAAACAAACCTCTGACAAAGTAGTAGCTGTGTAAGAAATTTGTGGAATCTGAAGAT 120
Db 198 TGCAGAAACAAACCTCTGACAAAGTAGTAGCTGTGTAAGAAATTTGTGGAATCTGAAGAT 257
QY 121 GATCCTGTGTTAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAAACAT 180
Db 258 GATCCTGTGTTAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAAACAT 317
QY 181 CCAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAAGGAAATTTGTGGAATCTGAAGAT 240
Db 318 CCAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAAGGAAATTTGTGGAATCTGAAGAT 377
QY 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT 300
Db 378 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT 437
QY 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCATATACAT 360
Db 438 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCATATACAT 497
QY 361 AACTGTATTACAGAGATATATAACCTGAAATATTTCTAATACTAAGCAAGGAATAATC 420
Db 498 AACTGTATTACAGAGATATATAACCTGAAATATTTCTAATACTAAGCAAGGAATAATC 557

QY 421 AAGATTGTGACTTCGGGTTTGCACAAATTCTG----- 453
Db 558 AAGATTGTGACTTCGGGTTTGCACAAATTCTGAGTTGGACTTCATCTTCTCTGGTGCC 617
QY 454 ----- 453
Db 618 TCCTTGATTGGCTTAATAGTTGACCTTCTGAAATCTTTTCTGCCAATTCAGAGATTTT 677
QY 454 ----- 453
Db 678 CTCCTGGCTTGGATCCATGTCTGACACAGTGTTCACCATGGGGCCAGGCTCATCTCGA 737
QY 454 ----- 453
Db 738 ACTTCTGGCCTCAAAGTGATCCTTCCACCTCGGCTCCCAAAGTGCTGGATTGCAAGTGTG 797
QY 454 ----- 453
Db 798 AGCCACCGTGCCCGAGCCAGATTTTTCAAAACAATAACTACTGAGAGCTCACAAGATTGTT 857
QY 454 -----ATTCCAGGAGATGCCTACACC 474
Db 858 TTAGTGGGAACACAAATTTCCGAACAAATTTCTTGAGAACGCATTCAGGAGATGCCTACACC 917
QY 475 GATTATGTAGCTACGAGATGTGTACCGAGCTCCTCGAACTTCTTGTGGGAGATACTCAGTAT 534
Db 918 GATTATGTAGCTACGAGATGTGTACCGAGCTCCTCGAACTTCTTGTGGGAGATACTCAGTAT 977
QY 535 GGTTCTTCAGTCGATATATGGCTATTGGTTGTGTTTTTGCAGAGCTCCTGACAGGCCAG 594
Db 978 GGTTCTTCAGTCGATATATGGCTATTGGTTGTGTTTTTGCAGAGCTCCTGACAGGCCAG 1037
QY 595 CCACTGTGGCCTGGAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGTA 654
Db 1038 CCACTGTGGCCTGGAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGTA 1097
QY 655 GAGACGGGGTTTCGCCATGTGTGACCGAGGTGTTCTCGAACTCTTGACGTCAAGTGATCCA 714
Db 1098 GAGACGGGGTTTCGCCATGTGTGACCGAGGTGTTCTCGAACTCTTGACGTCAAGTGATCCA 1157
QY 715 CCTGCCGTAGCCTCTCAAAGTCTTGAATTTACAGGAAAAATTAATCCCAAGACATCAATCA 774
Db 1158 CCTGCCGTAGCCTCTCAAAGTCTTGAATTTACAGGAAAAATTAATCCCAAGACATCAATCA 1217
QY 775 ATCTTTTAAAGTAACGGGGTTTTCATGGCATCAGTATACCTGAGCCAGAGACATGGAA 834
Db 1218 ATCTTTTAAAGTAACGGGGTTTTCATGGCATCAGTATACCTGAGCCAGAGACATGGAA 1277
QY 835 ACTCTTGAGGAAAAAGTTCTCAGATGTTTCATCTCTGTGGCTCTGAACTTTCATGAGGGGTGT 894
Db 1278 ACTCTTGAGGAAAAAGTTCTCAGATGTTTCATCTCTGTGGCTCTGAACTTTCATGAGGGGTGT 1337
QY 895 CTGAAGATGAATCCAGATGACAGATTAACTGTTCCCAACTCCTGGAGAGCTCCTACTTT 954
Db 1338 CTGAAGATGAATCCAGATGACAGATTAACTGTTCCCAACTCCTGGAGAGCTCCTACTTT 1397
QY 955 GATTCTTTCAAGAGGCCCAAAATTAAAGAAAAAGCACGTAATGAAGGAAGAAACAGAGA 1014
Db 1398 GATTCTTTCAAGAGGCCCAAAATTAAAGAAAAAGCACGTAATGAAGGAAGAAACAGAGA 1457
QY 1015 CGCCAACAGGT 1025
Db 1458 CGCCAACAGGT 1468

RESULT 8
AX286069
LOCUS AX286069 687 bp DNA linear PAT 20-NOV-2001
DEFINITION Sequence 3 from Patent WO0179488.
ACCESSION AX286069
VERSION AX286069.1 GI:17045995
KEYWORDS Homo sapiens (human)
SOURCE


```
RESULT 10
AX056404
LOCUS AX056404 911 bp DNA linear PAT 13-JAN-2001
DEFINITION Sequence 48 from Patent WO0373469.
ACCESSION AX056404
VERSION AX056404.1 GI:12229111
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Plozman,G.D., Martinez,R., Whyte,D. and Sudersanam,S.
TITLE Protein kinases
JOURNAL Patent: WO 0073469-A 48 07-DEC-2000;
Sugen, Inc. (US)
FEATURES
source
1..911
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 293 a 182 c 178 g 258 t
ORIGIN
Query Match 51.6%; Score 537; DB 6; Length 911;
Best Local Similarity 87.0%; Pred. No. 7.2e-120;
Matches 643; Conservative 0; Mismatches 0; Indels 96; Gaps 1;
QY 285 AAATGGAGTTGCTGATGAGTGTATCAAAAGCGGTATATGGCAACACTTCAGCTCTTAA 344
Db 1 AAATGGAGTTGCTGATGAGTGTATCAAAAGCGGTATATGGCAACACTTCAGCTCTTAA 60
QY 345 TTTCGTCTATATACATAACTGTATTACAGAGATATATAAACCTGAAATATTCTAATAAC 404
Db 61 TTTCGTCTATATACATAACTGTATTACAGAGATATATAAACCTGAAATATTCTAATAAC 120
QY 405 TAAGCAAGGAATAATCAAGATTGTGACTTCGGGTTTGACAAATTCGATTCAGGAGA 464
Db 121 TAAGCAAGGAATAATCAAGATTGTGACTTCGGGTTTGACAAATTCGATTCAGGAGA 180
QY 465 TGCTACACCGATTATGTAGTACGAGATGGTACCGAGCTCTGAACTTCTTGGGAGA 524
Db 181 TGCTACACCGATTATGTAGTACGAGATGGTACCGAGCTCTGAACTTCTTGGGAGA 240
QY 525 TACTCAGTATGGTCTTCAGTCGATATATGGGCTATTGGTGTGTTTGCAGAGCTCT 584
Db 241 TACTCAGTATGGTCTTCAGTCGATATATGGGCTATTGGTGTGTTTGCAGAGCTCT 300
QY 585 GACAGGCCAGCCACTGTGGCTGGAAATCAGATGTGGACCAACTTTATCTGATAATCAG 644
Db 301 GACAGGCCAGCCACTGTGGCTGGAAATCAGATGTGGACCAACTTTATCTGATAATCAG 360
QY 645 AACACTAGTAGAGACGGGGTTTCGCCATTTTGACCGGCTGGTCTCGAACTCTTGACGTC 704
Db 361 AACACT- 366
QY 705 AAGTGATCCACCTGCCGTAGCCTCTCAAGTGTCTGGAATTACAGGAATTAATCCCAAG 764
Db 367 - 366
QY 765 ACATCAATCAATCTTTAAAGTAACGGGTTTTTCCATGGCATCAGTATACCTGAGCCAGA 824
Db 385 ACATCAATCAATCTTTAAAGTAACGGGTTTTTCCATGGCATCAGTATACCTGAGCCAGA 444
QY 825 AGACATGGAAGTCTTGAGGAAAGTCTCTCAGATGTCTGCTGCTGCTGAACTTCAT 884
Db 445 AGACATGGAAGTCTTGAGGAAAGTCTCTCAGATGTCTGCTGCTGCTGAACTTCAT 504
QY 885 GAAGGGGTCTCTGAAGATGAATCCAGATGACAGATTAACTGTTCCTCCAACTCTCGGAGAG 944
Db 505 GAAGGGGTCTCTGAAGATGAATCCAGATGACAGATTAACTGTTCCTCCAACTCTCGGAGAG 564
```

```
QY 945 CTCCTACTTTGATCTTTTCAAGAGGCCCAAAATTAAAGAAAAGCACGTATGAAGGAAG 1004
Db 565 CTCCTACTTTGATCTTTTCAAGAGGCCCAAAATTAAAGAAAAGCACGTATGAAGGAAG 624
QY 1005 AAACAGAGAGACGCCCAACAG 1023
Db 625 AAACAGAGAGACGCCCAACAG 643
RESULT 11
AX107712
LOCUS AX107712 561 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 1 from Patent WO0123579.
ACCESSION AX107712
VERSION AX107712.1 GI:13923197
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and Sands,A.T.
TITLE Human kinase proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0123579-A 1 05-APR-2001;
Lexicon Genetics Incorporated (US)
FEATURES
source
1..561
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 189 a 87 c 112 g 173 t
ORIGIN
Query Match 43.7%; Score 454.4; DB 6; Length 561;
Best Local Similarity 99.8%; Pred. No. 8.3e-100;
Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGAAGATATGAAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAAA 60
Db 1 ATGGAAGATATGAAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAAA 60
QY 61 TGCAGAAACAAAACCTCTGGACAAGTAGTAGCTGTAAAAAATTTGTGGAATCTGAAGAT 120
Db 61 TGCAGAAACAAAACCTCTGGACAAGTAGTAGCTGTAAAAAATTTGTGGAATCTGAAGAT 120
QY 121 GATCCTGTTGTTAAGAAATAGCATAAGAGAAATACGATGTTGAAGCAATTAACAAT 180
Db 121 GATCCTGTTGTTAAGAAATAGCATAAGAGAAATACGATGTTGAAGCAATTAACAAT 180
QY 181 CCAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAGGAAAATGCAATTAGTTTT 240
Db 181 CCAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAGGAAAATGCAATTAGTTTT 240
QY 241 GAATAGTGTGATCATACACTTTTAAATGAGCTGGAAAGAACCCAAATGGAGTTGCTGAT 300
Db 241 GAATAGTGTGATCATACACTTTTAAATGAGCTGGAAAGAACCCAAATGGAGTTGCTGAT 300
QY 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCATATACAT 360
Db 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCATATACAT 360
QY 361 AACTGTATTCACAGAGATATAAAACCTGAAAATATTCTAATAACTAAGCAAGGAATAATC 420
Db 361 AACTGTATTCACAGAGATATAAAACCTGAAAATATTCTAATAACTAAGCAAGGAATAATC 420
QY 421 AAGATTTGTGACTTCGGGTTTGACAAATTTCTGATT 456
Db 421 AAGATTTGTGACTTCGGGTTTGACAAATTTCTGATT 456
RESULT 12
```

AX107718
LOCUS AX107718 594 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 7 from Patent WO0123579.
ACCESSION AX107718
VERSION AX107718.1 GI:13923200
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE 1
AUTHORS Donohue, G., Turner, C.A., Nehls, M., Friedrich, G., Zambrowicz, B. and Sands, A.T.
TITLE Human kinase proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0123579-A 7 05-APR-2001;
Lexicon Genetics Incorporated (US)
FEATURES
source
1..594
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 203 a 94 c 115 g 182 t
ORIGIN
Query Match 43.7%; Score 454.4; DB 6; Length 594;
Best Local Similarity 99.8%; Pred. No. 8.3e-100;
Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGAAGAAGTATGAAAAATTAGTAAGACTGGAGAGAGGCTCTTATGGGGTGTATTCAA 60
DB 1 ATGGAAGAAGTATGAAAAATTAGTAAGACTGGAGAGAGGCTCTTATGGGGTGTATTCAA 60
QY 61 TGCAGAAACAAACCTCTGGACAAGTAGTAGCTGTTAAAAAATTGTGGAACTCTGAAGAT 120
DB 61 TGCAGAAACAAACCTCTGGACAAGTAGTAGCTGTTAAAAAATTGTGGAACTCTGAAGAT 120
QY 121 GATCCTGTGTTAAGAAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAACAT 180
DB 121 GATCCTGTGTTAAGAAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAACAT 180
QY 181 CCAATCTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAAGGAAATGCATTTAGTTTTT 240
DB 181 CCAATCTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAAGGAAATGCATTTAGTTTTT 240
QY 241 GAATACGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT 300
DB 241 GAATACGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT 300
QY 301 GGAGTGATCAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGTCAATATACAT 360
DB 301 GGAGTGATCAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGTCAATATACAT 360
QY 361 AACTGTATTACAGAGATATATAAACCTGAAAAATATTTCTAATAACTAAGCAAGGAATAATC 420
DB 361 AACTGTATTACAGAGATATATAAACCTGAAAAATATTTCTAATAACTAAGCAAGGAATAATC 420
QY 421 AAGATTGTGACTTCGGGTTTGACAAATTTCTGATT 456
DB 421 AAGATTGTGACTTCGGGTTTGACAAATTTCTGAGT 456
RESULT 13
AX056405
LOCUS AX056405 2615 bp DNA linear PAT 13-JAN-2001
DEFINITION Sequence 49 from Patent WO0073469.
ACCESSION AX056405
VERSION AX056405.1 GI:12229112
KEYWORDS Murinae gen. sp.
SOURCE Murinae gen. sp.
ORGANISM Murinae gen. sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae.
1

AUTHORS Plowman, G.D., Martinez, R., Whyte, D. and Sudersanam, S.
TITLE Protein kinases
JOURNAL Patent: WO 0073469-A 49 07-DEC-2000;
Sugen, Inc. (US)
FEATURES
Location/Qualifiers
source
1..2615
/organism="Murinae gen. sp."
/mol_type="genomic DNA"
/db_xref="taxon:39108"
BASE COUNT 742 a 580 c 585 g 708 t
ORIGIN
Query Match 43.5%; Score 453.2; DB 6; Length 2615;
Best Local Similarity 73.1%; Pred. No. 1.6e-99;
Matches 658; Conservative 0; Mismatches 143; Indels 99; Gaps 2;
QY 135 GAAAAATAGCACTAAGAGAAAAATACGTATGTTGAAGCAATTAATAACATCCAAATCTTGTA 194
DB 1 GAAAAATAGCCCTGCGGGNAATCCGTATGCTGAAG---TTGAACACACCCAAACCTCGTGA 57
QY 195 CCTCATCGAGGTGTTTCAGGAGAAAAAGAAAAATGCATTTAGTTTTTGAATACCTGATCA 254
DB 58 CCTCATCGAGGTGTTTCAGGAGAAAAAGAAAAATGCATCTAGTTTTTGAGTACTGATCA 117
QY 255 TACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGATGGAGTATCAAAAG 314
DB 118 CACACTGTTAAACGAGCTGGAGAGAAACCCAAACGGAGTTTCTGATGGAGTATTAAG 177
QY 315 CGTATTATGSCAAACACTTCAAGCTCTTAATTTCTGTATATACATACTGTATTCACAG 374
DB 178 TGTGCTATGGCAAAACCTTCAAGCCCTTAACCTTCTGTCAAGACACAATTTGATTCATCG 237
QY 375 AGATATAAAACCTGAAATATTTCTAATAACTAAGCAAGGAATAATCAAGATTTTGACTT 434
DB 238 GGATGTAACCTGAAACATCTCTAATAACCAAGCAAGGGATGATAAAGATTTTGACTT 297
QY 435 CGGTTTTGCACAAATCTGATTTCCAGGAGATGCCTACACCGGATTTATGATCTACGAGATG 494
DB 298 TGSATTTGCACGAATCTAATTTCCAGGAGACGCCTACACAGACTATGTTGCCACCGGTG 357
QY 495 GTACCGAGCTCTGAACTTCTTGTGGAGATACTCAGTATGGTTCTTCAGTCGATATATG 554
DB 358 GTACCGAGCCCCGAACTTCTCGTGGAGACACGAAGTACSGTCTCTGTAGACGTGTG 417
QY 555 GGCTATTGGTTGTGTTTTTGCAGAGCTCCTGACAGGCCAGCCACTGTGGCCTGGAAATC 614
DB 418 GGCCGTGGGCTGTGTTTTTGCAGAGCTCTGACGGGTGACGCCACTCTGGCCGGGAAATC 477
QY 615 AGATGTGGACCAACTTTATCTGATATATCAGAAACACTAGTAGACGGGTTTCGCCATGT 674
DB 478 CGACGTGGACCCAGCTTTACCTGATCATCAGGACGTTG----- 514
QY 675 TGACCGAGCTGGTCTCGAACTCTTGACCTCAAGTGATCCACCTGCCGTAGCCTCTCAAG 734
DB 515 ----- 514
QY 735 TGCTGGAATTACAGGAAAAATTAATCCCAAGACATCAATCAATCTTTAAAAAGTAACGGTT 794
DB 515 -----GGGAAGCTGATTCGAAGACACCAGTCTATCTTTAGGAGTAACCAAGTT 561
QY 795 TTTCCATGGCATCAGTATACCTGAGCCAGAAAGACATGGAACCTCTTGAGGAAAAAGTTCTC 854
DB 562 TTTCCGCGGCATCAGCATACCTGAAACAGAGGACATGGAGACTCTTGAAGAAAAAATTCTC 621
QY 855 AGATGTTTCATCCTGTGGCTCTGAACTTCATGAAGGGGTGTCGAAGATGAATCCAGATGA 914
DB 622 AAATGTTTCAGCCTGTGGCTTTAAGTTTCATGAAGGGATGCCTGAAGATGAATCCTGATGA 681
QY 915 CAGATTAACTGTTCCCAACTCCTGGAGAGCTCCTACTTTGATTTCTTTCAAGAGGCCCA 974
DB 682 GAGGCTGACCTGTGCCCAGCTGCTGGACAGTGCCTACTTTGAGTCTTTTCAAGAGGATCA 741
QY 975 AATTAAAGAAAAAGCACGTAATGAAGGAAGAAACAGAGACGCCCAACAGGTACTTCCGCT 1034

Db 742 AATGAAAGAAAGCCCGAGTGAGGGAGAGCCGAGGCGCCAGAGAAATCAACTGCT 801

RESULT 14
HSSTHPKB
LOCUS HSSTHPKB 1363 bp mRNA linear PRI 06-FEB-1997
DEFINITION H.sapiens mRNA KIA1ALRE for serine/threonine protein kinase.
ACCESSION X66358
VERSION X66358.1 GI:36614
KEYWORDS serine/threonine protein kinase.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1363)
AUTHORS Meyerson,M., Enders,G.H., Wu,C.L., Su,L.K., Gorka,C., Nelson,C.,
Harlow,E. and Tsai,L.H.
TITLE A family of human cdc2-related protein kinases
JOURNAL EMBO J 11 (8), 2909-2917 (1992)
MEDLINE 92347325
PUBMED 1639063
REFERENCE 2 (bases 1 to 1363)
AUTHORS Meyerson,M.L.
TITLE Direct Submission
JOURNAL Submitted (12-MAY-1992) M.L. Meyerson, Massachusetts General
Hospital, Cancer Center, Bldg 149, 13th Street, Charlestown MA
02129, USA
FEATURES
source Location/Qualifiers
1. 1363
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="HeLa"
/cell_type="HeLa"
/clone_lib="HeLa"
214..1290
/codon_start=1
/product="serine/threonine protein kinase"
/protein_id="CAA47032.1"
/db_xref="GI:36615"
/db_xref="SWISS-PROT:Q00532"
/translation="YMEKYEKIGKIGEGSYGVVFKCRNRDTGQIVAIKKFLESEDDPV
IKKIALREIRMLKQLKHPNLVNLLEVFRRKRRLHLVFEYCDHTVLHEDRYQGVPEH
LVKSIWTQTLQAVNFCCHKNCIHRDVKNENILITKHSVIKLCDFGFARLLTGPSDYIT
DYVATRWYRSPPELLVGTQYGPVVDVWAGCVFAELLSGVPLWPGKSDVDQYLIRKT
LGDLPFRHQVSTNQYFSGVKIPDPEDYEPLELKFNPISYPALGLLKGCLHMDPTER
LTCEQLLHPYPENIREIEDLAKHEHDKPTKRLKSRKHHCFTETSKLQYLPQLTGSS
ILPALDNKKYYCDTKLNYRFPNI"
join(214..669,X66359.1:1..54,670..1290)
/note="KKIALRE coding sequence with 54 bp insertion"
/pseudo
/codon_start=1
/label=sthpkln CDS
/product="serine/threonine protein kinase"
/db_xref="PID:e49306"
BASE COUNT 412 a 308 c 309 g 334 t
ORIGIN

Query Match 33.0%; Score 343.8; DB 3; Length 1363;
Best Local Similarity 71.2%; Pred. No. 5.9e-73;
Matches 469; Conservative 0; Mismatches 187; Indels 3; Gaps 1;

QY 1 ATGGAAAGAAATAGTAAGACTGGAGAGGGTCTTATGGGGTTGATTCAA 60
|||||
Db 217 ATGGAGAGATGAAAAAATGGGAAATGGAGAGGATCCTATGGAGTTGTTTCAA 276
|||||
QY 61 TGCAGAACAAACCTCTGGACAGTAGTAGCTGTTAAAAAATTTGTGGAACTCTGAAGAT 120
|||||
Db 277 TGTAGAACAGGGACACGGGTGAGATTGTGGCCATCAAGAAGTTTCTGGAAATCAGAAGAT 336
|||||
QY 121 GATCCTGTGTTAAGAAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAACAT 180
|||||

Db 337 GACCCCTGTCAATAAGAAATTCGCCCTTCGGGAAATCCGAATGCTCAAGCAACTCAAGCAT 396

QY 181 CCAAAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAAGGAAATGCAATTTAGTTTTT 240
|||||
Db 397 CCCAACCTTGTTAACCTCCTCGAAGTCTTCAGGAGGAAACGGAGGCTTCACCTGCTGTTT 456
|||||

QY 241 GAATAGTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT 300
|||||
Db 457 GAATATTGTGACCAACACAGTCTCCATGAGTTGGACAGATACCAAGAGGGGTACCAAGAA 516
|||||

QY 301 GGAGTGATCAAAAGCGTATTATGGCAACACCTTCAGGCTCTTAATTTCTGTGATATACAT 360
|||||
Db 517 CATCTCGTGAAGAGCATAACTTGGCAGACACTGCAAGCTGTAATTTTGGCATAAACAC 576
|||||

QY 361 AACTGTATTCAAGAGATATATAACCTGAAAAATATTCTAATAACTAAGCAAGAAATAATC 420
|||||
Db 577 AATTGCATACATAGAGACGCTGAAGCCAGAAAAATATCCTCATCACGAAACATTCGCTGATT 636
|||||

QY 421 AAGATTTGTGACTTCGGGTTTGGCAAAATTTCTGATT---CCAGGAGATGCCCTACACCGAT 477
|||||
Db 637 AAGCTTTGTGACTTTGGATTTCGCTCGGCTTTTGAATGACCGAGTGACTACTATACAGAC 696
|||||

QY 478 TATGTAGCTACGAGATGGTACCGAGCTCCTGAACCTTCTTGTGGAGATACTCAGTATGGT 537
|||||
Db 697 TACGTGGCTACACAGGTGGTACCGCTCCCTGAGCTGCTGGTGGGACACGAGTACGGC 756
|||||

QY 538 TCTTCAGTGCATATATGGGCTATGTTGTTGTTTGTGTTTTCAGAGAGCTCCTGACAGGCCAGCCA 597
|||||
Db 757 CCCCCGGTGGATGTTTGGGCAATTTGGCTGTGCTTTTGTGAGCTGCTGTGAGAGTGCTCT 816
|||||

QY 598 CTGTGGCCTGGAAAAATCAGATGTGACCAACCTTTATCTGATAATCAGAAACACTAGTAGA 656
|||||
Db 817 CTGTGGCCAGGAAATCGGATGTGGATCAGCTGTATCTGATTAGGAAGACCTTGGGGGA 875
|||||

RESULT 15
AX399975
LOCUS AX399975 1612 bp DNA linear PAT 06-JUN-2002
DEFINITION Sequence 146 from Patent WO0218424.
ACCESSION AX399975
VERSION AX399975.1 GI:21336301
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Tang,Y.T., Asundi,V., Zhou,P., Xue,A.J., Ren,F., Zhang,J.,
Wang,J.R., Zhao,Q.A., Wang,D., Liu,C., Drmanac,R.T. and Wehrman,T.
TITLE Nucleic acids and polypeptides
JOURNAL Patent: WO 0218424-A 146 07-MAR-2002;
HYSEQ, INC. (US)
FEATURES
source Location/Qualifiers
1. 1612
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
463..1539
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD33419.1"
/db_xref="GI:21336302"
/translation="MMEKYEKIGKIGEGSYGVVFKCRNRDTGQIVAIKKFLESEDDPV
IKKIALREIRMLKQLKHPNLVNLLEVFRRKRRLHLVFEYCDHTVLHEDRYQGVPEH
LVKSIWTQTLQAVNFCCHKNCIHRDVKNENILITKHSVIKLCDFGFARLLTGPSDYIT
DYVATRWYRSPPELLVGTQYGPVVDVWAGCVFAELLSGVPLWPGKSDVDQYLIRKT
LGDLPFRHQVSTNQYFSGVKIPDPEDMEPLELKFNPISYPALGLLKGCLHMDPTER
LTCEQLLHPYPENIREIEDLAKHEHDKPTKRLKSRKHHCFTETSKLQYLPQLTGSS
ILPALDNKKYYCDTKLNYRFPNI"
BASE COUNT 476 a 380 c 361 g 395 t
ORIGIN

Query Match 33.0%; Score 343.8; DB 6; Length 1612;

Best Local Similarity 71.2%; Pred. No. 5.9e-73;		Matches 469; Conservative 0; Mismatches 187; Indels 3; Gaps 1;	
QY	1	ATGGAAAAGTATGAAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGGTTGTATTCAAA	60
Db	466	ATGGAGAAGTATGAAAAATTGGGAAAAATTGGAGAGGATCCTATGGAGTTGTTTCAAA	525
QY	61	TGCAGAAACAAAACCTCTGGACAAGTAGTAGCTGTAAAAAAATTTGIGGAATCTGAAGAT	120
Db	526	TGTAGAAACAGGGACACCGGTGAGATGTGGCCATCAAGAAGTTTCTGGAATCAGNAGAT	585
QY	121	GATCCTGTTGTTAAAGAAAATAGCACTAAAGAGAAATACGTATGTTGAAGCAATTAAAAACAT	180
Db	586	GACCCGTGCTAAAGAAAATTTGCCCTTCGGGAAATCCGAATGCTCAAGCAACTCAAGCAT	645
QY	181	CCAAATCTTGTGAACCTTCATCGAGGTGTTTCAGGAGAAAAAGGAAAAATGCAATTTAGTTTTT	240
Db	646	CCCAACCTTGTTAAACCTCCCTGGAAAGTCTTCAGGAGGAAACGGAGGCTTCACCTGSGTGT	705
QY	241	GAATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCRAATGGAGTTGCTGAT	300
Db	706	GAATATTGTGACCACACAGTCTCCATGAGTTGGACAGATACCAAAAGAGGGGTACCAAGAA	765
QY	301	GGAGTGATCAAAAGCGTATTATGGCAAAACACTTCAGCTCTTAATTTCTGTGCATATACAT	360
Db	766	CATCTCGTGAAGAGCCTAAGTGGCAGACACTGCAAGCTGTAAATTTTGGCATAAACAC	825
QY	361	AACTGTATTCCAGAGATATAAAACCTGAAAATATTCTAATAACTAAGCAAGGAATATC	420
Db	826	AATTGCATACATAGAGCGTGAAGCCAGAAAAATATCCTCATCACGNAACATTCCTGATT	885
QY	421	AAGATTGTGACTTCGGGTTTGCACAAAATTCIGATT---CCAGGAGATGCTACACCGAT	477
Db	886	AAGCTTTGTGACTTTGGATTGGCTGGCTTTTGACTGGAACCGAGTGACTACTATACAGAC	945
QY	478	TATGTAGCTACGAGATGGTACCGAGCTCCTGAACTTCTGTGGGAGATACATCAGTATGGT	537
Db	946	TACGTGGCTACCAAGGTGGTACCCCTCCCTGAGCTGCTGGTGGGGACACGCAAGTACGGC	1005
QY	538	TCTTCAGTCGATATATGGGCTATTGGTGTGTTTTTGCAGAGCTCCTGACAGGCCAGCCA	597
Db	1006	CCCCCGGTGGATGTTTGGGCAA-TGGCTGTGTCTTTGCTGAGCTGCTGTGAGGAGTGCCT	1065
QY	598	CTGTGGCCTGGAAAAATCAGATGTGGACCAACTTTTATCTGATAATCAGAACACTAGTAGA	656
Db	1066	CTGTGGCCAGGAAAAATCGGATGTGGATCAGCTGTATCTGATTAGGAAGACCTTGGGGGA	1124

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OK: protein - protein search, using sw model

Run on: November 13, 2003, 14:12:21 ; Search time 39.8369 Seconds
(without alignments)
1382.592 Million cell updates/sec

Title: US-09-671-050-10
Perfect score: 1820
Sequence: 1 MEKYEKLAKTGEGSYGVVFK.....RKARNEGRRRRQQVLP LKS 347

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq 19Jun03:*

1:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
23:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*
24:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1820	100.0	347	AAE00494	Human kinase #5.
2	1796	98.7	356	AAE00491	Human kinase #2.
3	1636	89.9	315	AAE00495	Human kinase #6.
4	2613	88.6	342	AAE19152	Human kinase polyp
5	1612	88.6	324	AAE00492	Human kinase #3.
6	1593	87.5	360	AAU03525	Human protein kina
7	1270.5	69.8	296	AAE00493	Novel protein kina
8	1162.5	63.9	358	ABP62954	Human polypeptide
9	1134	62.3	228	AAG78547	Human kinase 14257

10	1123	61.7	247	22	AAE00492	Novel protein kina
11	1001.5	55.0	392	22	ABE63118	Drosophila melanog
12	948	52.1	197	24	ABU11689	Human MDDT polypep
13	941.5	51.7	566	21	AAE00494	Rabbit KKIAMRE kin
14	911.5	50.1	362	24	ABP96087	Human protein kina
15	791	43.5	187	22	AAE00490	Human kinase #1.
16	791	43.5	198	22	AAE00493	Human kinase #4.
17	758.5	41.7	455	21	AAE00492	Human lost in leuk
18	758.5	41.7	591	22	AAU03524	Human protein kina
19	737.5	40.5	457	21	AAE00492	Rat lost in leukae
20	737.5	40.5	505	21	AAE00492	Rat lost in leukae
21	712.5	39.1	154	24	ABU05290	Human diagnostics
22	666	36.6	534	22	AAE00494	Novel protein kina
23	569	31.3	224	22	ABG06142	Novel human diagno
24	565	31.0	333	22	AAE00492	Human polypeptide
25	563	30.9	544	19	AAE00492	CDK2-green fluores
26	558	30.7	352	22	AAE00492	Human polypeptide
27	557	30.6	298	20	AAE00492	Homo sapiens Cdc2
28	557	30.6	544	19	AAE00492	Green fluorescent
29	555	30.5	297	23	ABE79004	Human 10Ckshs1 pro
30	551.5	30.3	294	20	AAE00492	Oryza sativa Cdc2
31	551.5	30.3	294	21	AAE00492	Arabidopsis thalia
32	551.5	30.3	294	21	AAE00492	Arabidopsis thalia
33	551.5	30.3	250	21	AAE00492	Arabidopsis thalia
34	547.5	30.1	270	20	AAE00492	Human protein kina
35	546.5	30.0	294	21	AAE00492	Zea mays protein f
36	545	29.9	297	20	AAE00492	Homo sapiens Cdc2
37	545	29.9	297	22	AAE00492	Amino acid sequenc
38	545	29.9	297	22	AAE00492	Amino acid sequenc
39	545	29.9	297	24	ABE47405	Breast cancer asso
40	545	29.9	297	24	AAE00492	Human CDC2 protein
41	545	29.9	297	24	ABU56522	Lung cancer-associ
42	541.5	29.8	294	21	AAE00492	Zea mays protein f
43	541	29.7	297	20	AAE00492	Rattus norvegicus
44	541	29.7	298	21	AAE00492	Cyclin-dependent k
45	540.5	29.7	297	22	ABE61462	Drosophila melanog

ALIGNMENTS

RESULT 1
AAE00494
ID AAE00494 standard; Protein; 347 AA.
AC AAE00494;
XX
XX
XX 19-JUN-2001 (first entry)
XX
XX Human kinase #5.
XX
XX Human; kinase; gene therapy; bioreactor; mental disorder;
XX biological disorder.
XX Homo sapiens.
XX
XX WO200123579-A1.
XX
XX 05-APR-2001.
XX
XX 27-SEP-2000; 2000WO-US26621.
XX
XX 28-SEP-1999; 99US-0156511.
XX
XX (LEXI-) LEXICON GENETICS INC.
XX
XX Dorohc G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
XX WPI; 2001-266166/27.
XX N-PSDB; AAD03816.
XX
XX New isolated human kinase polynucleotide useful for generating
XX antibodies, as reagents in diagnostic assays and for screening for

PT compounds useful for treating mental, biological or medical diseases -

XX Claim 2; Page 32-33; 38pp; English.

XX The present sequence is novel human protein.

CC (NHP) known as human kinase. The human kinases share structural

CC similarity with animal kinases, more particularly serine or

CC threonine protein kinases. Human kinase cDNA is useful for the

CC detection of mutant human kinase for the diagnosis of disease,

CC and also as a therapeutic. It is useful for screening drugs

CC effective in the treatment of symptomatic or phenotypic

CC manifestations perturbing the normal function of NHP in the

CC body. The NHP nucleotide sequences are useful for generation of

CC antibodies, as reagents in diagnostic assays, for the

CC identification of other cellular gene products related to human

CC kinases, and as reagents in assays for screening compounds that

CC are useful for treating mental, biological or medical disorders.

CC NHP oligonucleotides are used as probes. The labelled NHP probes

CC are useful for screening human genomic library for identifying

CC polymorphisms and as primers in amplification assays to detect

CC mutations within the exons, introns and splice sites that can

CC be used in diagnostics and pharmacogenomics. Nucleotide construct

CC encoding NHP products are used to genetically engineer cells

CC in vivo that functions as bioreactors in the body delivering a

CC continuous supply of NHP to the body. Nucleotide constructs

CC encoding functional NHPs are used in gene therapy for the

CC modulation of NHP expression.

XX

SQ Sequence 347 AA;

Query Match 100.0%; Score 1820; DB 22; Length 347;

Best Local Similarity 100.0%; Pred. No. 1.7e-186;

Matches 347; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEKYEKLAKTGECSYGVVFKCRNKTSGQVAVKKFVESEDDPVVKIALREIRMLKQLKH 60

DB 1 MEKYEKLAKTGECSYGVVFKCRNKTSGQVAVKKFVESEDDPVVKIALREIRMLKQLKH 60

QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEPNPVGADGVKSVLWQTLQALNFCIH 120

DB 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEPNPVGADGVKSVLWQTLQALNFCIH 120

QY 121 NCIHRIKPNILITKQGIKICDFGAQILIPGDYTDYVATRWYRAPELLVGDYQGS 180

DB 121 NCIHRIKPNILITKQGIKICDFGAQILIPGDYTDYVATRWYRAPELLVGDYQGS 180

QY 161 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIRTLVETGFRHVDQAGLELLTSSDPPA 240

DB 161 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIRTLVETGFRHVDQAGLELLTSSDPPA 240

QY 241 VASQSAGITGKLIPRHQSIFKSNFGFFHGISIPEPEDMETLEEKFSVHPVALNFMKGCLK 300

DB 241 VASQSAGITGKLIPRHQSIFKSNFGFFHGISIPEPEDMETLEEKFSVHPVALNFMKGCLK 300

QY 301 MNPDRLTCSQLLESSYFDSFQEAQIKRKARNEGRRRRQOVLPKLS 347

DB 301 MNPDRLTCSQLLESSYFDSFQEAQIKRKARNEGRRRRQOVLPKLS 347

RESULT 2

AAE00491

ID AAE00491 standard; Protein; 356 AA.

XX

AC AAE00491;

XX

DT 19-JUN-2001 (first entry)

XX

DE Human kinase #2.

XX

KW Human; kinase; gene therapy; bioreactor; mental disorder;

KW biological disorder.

XX

OS Homo sapiens.

XX WO200123579-A1.

XX CS-APR-2001.

XX 27-SEP-2000; 2000WO-US26621.

XX 28-SEP-1999; 99US-0156511.

PA (LEXI-) LEXICON GENETICS INC.

XX

PI Donoho G, Turner CA, Nehls X, Friedrich G, Zambrowicz B, Sands AT;

XX WPI; 2001-266166/27.

DR N-PSDB; AAD03813.

XX

PT New isolated human kinase polynucleotide useful for generating

PT antibodies, as reagents in diagnostic assays and for screening for

PT compounds useful for treating mental, biological or medical diseases -

XX

PS Disclosure; Page 28-29; 38pp; English.

XX

CC The present sequence is novel human protein

CC (NHP) known as human kinase. The human kinases share structural

CC similarity with animal kinases, more particularly serine or

CC threonine protein kinases. Human kinase cDNA is useful for the

CC detection of mutant human kinase for the diagnosis of disease,

CC and also as a therapeutic. It is useful for screening drugs

CC effective in the treatment of symptomatic or phenotypic

CC manifestations perturbing the normal function of NHP in the

CC body. The NHP nucleotide sequences are useful for generation of

CC antibodies, as reagents in diagnostic assays, for the

CC identification of other cellular gene products related to human

CC kinases, and as reagents in assays for screening compounds that

CC are useful for treating mental, biological or medical disorders.

CC NHP oligonucleotides are used as probes. The labelled NHP probes

CC are useful for screening human genomic library for identifying

CC polymorphisms and as primers in amplification assays to detect

CC mutations within the exons, introns and splice sites that can

CC be used in diagnostics and pharmacogenomics. Nucleotide construct

CC encoding NHP products are used to genetically engineer cells

CC in vivo that functions as bioreactors in the body delivering a

CC continuous supply of NHP to the body. Nucleotide constructs

CC encoding functional NHPs are used in gene therapy for the

CC modulation of NHP expression.

XX

SQ Sequence 356 AA;

Query Match 98.7%; Score 1796; DB 22; Length 356;

Best Local Similarity 100.0%; Pred. No. 6.6e-184;

Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEKYEKLAKTGECSYGVVFKCRNKTSGQVAVKKFVESEDDPVVKIALREIRMLKQLKH 60

DB 1 MEKYEKLAKTGECSYGVVFKCRNKTSGQVAVKKFVESEDDPVVKIALREIRMLKQLKH 60

QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEPNPVGADGVKSVLWQTLQALNFCIH 120

DB 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEPNPVGADGVKSVLWQTLQALNFCIH 120

QY 121 NCIHRIKPNILITKQGIKICDFGAQILIPGDYTDYVATRWYRAPELLVGDYQGS 180

DB 121 NCIHRIKPNILITKQGIKICDFGAQILIPGDYTDYVATRWYRAPELLVGDYQGS 180

QY 161 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIRTLVETGFRHVDQAGLELLTSSDPPA 240

DB 161 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIRTLVETGFRHVDQAGLELLTSSDPPA 240

QY 241 VASQSAGITGKLIPRHQSIFKSNFGFFHGISIPEPEDMETLEEKFSVHPVALNFMKGCLK 300

DB 241 VASQSAGITGKLIPRHQSIFKSNFGFFHGISIPEPEDMETLEEKFSVHPVALNFMKGCLK 300

QY 301 MNPDRLTCSQLLESSYFDSFQEAQIKRKARNEGRRRRQOV 342

Db 301 MNPDRLTCSQLLESSYFDSFQEAQIKRKARNEGRRRRQV 342

RESULT 3
AAE00495
ID AAE00495 standard; Protein; 315 AA.
XX
AC AAE00495;
XX
DT 19-JUN-2001 (first entry)
XX
DE Human kinase #6.
XX
KW Human; kinase; gene therapy; bioreactor; mental disorder;
KW biological disorder.

XX Homo sapiens.
OS
XX WPI; 2001-266166/27.
XX N-PSDB; AAD03817.
XX
XX New isolated human kinase polynucleotide useful for generating
PT antibodies, as reagents in diagnostic assays and for screening for
PT compounds useful for treating mental, biological or medical diseases
XX
PS Claim 3; Page 33-34; 38pp; English.
XX

CC The present sequence is novel human protein
CC (NHP) known as human kinase. The human kinases share structural
CC similarity with animal kinases, more particularly serine or
CC threonine protein kinases. Human kinase cDNA is useful for the
CC detection of mutant human kinase for the diagnosis of disease,
CC and also as a therapeutic. It is useful for screening drugs
CC effective in the treatment of symptomatic or phenotypic
CC manifestations perturbing the normal function of NHP in the
CC body. The NHP nucleotide sequences are useful for generation of
CC antibodies, as reagents in diagnostic assays, for the
CC identification of other cellular gene products related to human
CC kinases, and as reagents in assays for screening compounds that
CC are useful for treating mental, biological or medical disorders.
CC NHP oligonucleotides are used as probes. The labelled NHP probes
CC are useful for screening human genomic library for identifying
CC polymorphisms and as primers in amplification assays to detect
CC mutations within the exons, introns and splice sites that can
CC be used in diagnostics and pharmacogenomics. Nucleotide construct
CC encoding NHP products are used to genetically engineer cells
CC in vivo that functions as bioreactors in the body delivering a
CC continuous supply of NHP to the body. Nucleotide constructs
CC encoding functional NHPs are used in gene therapy for the
CC modulation of NHP expression.

XX Sequence 315 AA;
SQ
Query Match 89.9%; Score 1636; DB 22; Length 315;
Best Local Similarity 90.8%; Pred. No. 8.6e-167;
Matches 315; Conservative 0; Mismatches 0; Indels 32; Gaps 1;

Qy 1 MEKYEKLAKTGSGYGWVFKCRNKTSGQWVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MEKYEKLAKTGSGYGWVFKCRNKTSGQWVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60

Qy 61 PNLVNLIEVFRKRKQKHLVFEYCDHTLLNELEPNPGVADGVKSVLQTLQALNFCIH 120
Db 61 PNLVNLIEVFRKRKQKHLVFEYCDHTLLNELEPNPGVADGVKSVLQTLQALNFCIH 120
Qy 121 NCIHRIKPENILITKQGIKICDFGPAQILIPGDAYTDYVATRWYRAPELLVGDTOYGS 180
Db 121 NCIHRIKPENILITKQGIKICDFGPAQILIPGDAYTDYVATRWYRAPELLVGDTOYGS 180
Qy 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLVETGFRHVDQAGLELLTSSDPPA 240
Db 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTL----- 217
Qy 241 VASQSAGITGKLIPRHQSIFKSNFGFHGISIPEPEDMETLEKFSVDVHPVALNFMKGCLK 300
Db 218 -----GKLIPRHQSIFKSNFGFHGISIPEPEDMETLEKFSVDVHPVALNFMKGCLK 268
Qy 301 MNPDRLTCSQLLESSYFDSFQEAQIKRKARNEGRRRRQVPLKS 347
Db 269 MNPDRLTCSQLLESSYFDSFQEAQIKRKARNEGRRRRQVPLKS 315

RESULT 4
AAE19152
ID AAE19152 standard; Protein; 342 AA.
XX
AC AAE19152;
XX
DT 21-MAY-2002 (first entry)
XX
DE Human kinase polypeptide (PKIN-10).

XX Human; kinase polypeptide; PKIN-10; gene therapy; Addison's disease;
KW leukaemia; immune disorder; lymphoma; melanoma; developmental disorder;
KW acquired immunodeficiency syndrome; AIDS; allergy; anaemia; hypertension;
KW asthma; Crohn's disease; rheumatoid arthritis; bursitis; atherosclerosis;
KW cirrhosis; hepatitis; psoriasis; Cushing's syndrome; cytostatic; cancer;
KW cholestasis; cardiact; cardiovascular disorder; Niemann-Pick's disease;
KW lipid disorder; fatty liver; Gaucher's disease; myocardial infarction;
KW drug screening; transgenic animal; antiinflammatory; hepatotropic;
KW hypotensive; anti-HIV; enzyme.

XX Homo sapiens.
CS
XX
PH Key Location/Qualifiers
FT Domain 4..286
FT /note= "Eukaryotic protein kinase domain"

XX WO200208399-A2.
PN
XX
PD 31-JAN-2002.
XX
PF 20-JUL-2001; 2001WO-US23092.
XX
PR 21-JUL-2000; 2000US-220038P.
PR 28-JUL-2000; 2000US-222112P.
PR 04-AUG-2000; 2000US-222831P.
PR 11-AUG-2000; 2000US-224729P.
XX
PA (INCY-) INCYTE GENOMICS INC.
PA (THOR/) THORNTON M.

XX
PI Thornton M, Yue H, Khan FA, Gururajan R, Hafalia AJA, Walia NK;
PI Patterson C, Ramkumar J, Gandhi AR, Policky JL, Baughn MR;
PI Tribouley CM, Bandman O, Nguyen DB, Lu Y, Burford N, Lal P;
PI Ding L, Yao MG, Elliott VS, Recipon SA, Kearney L, Lu DAM;
PI Greenwald SR, Tang YT, Xu Y, Walsh RT, Gietzen KJ, Yang J;
PI Hillman JL;
XX
DR WPI; 2002-206083/26.
XX N-PSDB; AAD30557.
PT New human kinase polypeptide, useful in diagnosis, prevention and

PT treatment of cancer, immune disorder, growth and developmental
PT disorder, cardiovascular disorder and lipid disorder -
XX
PS Claim 1; Page 147-148; 196pp; English.
XX

CC The present invention relates to an isolated human kinase polypeptide
CC (PKIN) or a biologically active/immunogenic fragment of PKIN. PKIN is
CC useful for diagnosing, treating and preventing cancer (e.g., leukaemia,
CC lymphoma, melanoma), an immune disorder (e.g., acquired immunodeficiency
CC syndrome (AIDS), Addison's disease, allergy, anaemia, asthma, Crohn's
CC disease, rheumatoid arthritis), a growth and developmental disorder (e.g.
CC bursitis, cirrhosis, hepatitis, psoriasis, Cushing's syndrome), a
CC cardiovascular disorder (e.g., atherosclerosis, hypertension, myocardial
CC infarction), and a lipid disorder (e.g., fatty liver, cholestasis,
CC Gaucher's disease, Niemann-Pick's disease). PKIN is useful in a number of
CC drug screening techniques and to analyse the proteome of a tissue or cell
CC type. PKIN is useful for creating knockin humanised animals or transgenic
CC animals to model human diseases, in somatic or germline gene therapy, to
CC generate a transcript image of a tissue or cell type, for detecting
CC differences in the chromosomal location due to translocation, inversion,
CC etc., among normal, carrier or affected individuals, and as hybridisation
CC probes for mapping naturally occurring genomic sequences. PKIN is useful
CC in southern or northern analysis, dot blot or other membrane-based
CC technologies, in PCR technologies, in dipstick, pin, multiformat enzyme
CC linked immunosorbent (ELISA)-like assays and in microarrays utilising
CC fluids or tissues from patients to detect altered PKIN expression. The
XX present sequence is human PKIN-10.

SQ Sequence 342 AA;

Query Match 88.6%; Score 1613; DB 23; Length 342;
Best Local Similarity 89.9%; Pred. No. 2.9e-164;
Matches 312; Conservative 1; Mismatches 0; Indels 34; Gaps 2;

QY 1 MEKYEKLAKTGEISYGVVFKCRNKTSGQVAVKFKVESEDDPVVKKIALREIRMLKQLKH 60
DB 1 MEKYEKLAKTGEISYGVVFKCRNKTSGQVAVKFKVESEDDPVVKKIALREIRMLKQLKH 60
QY 61 PNLVNLIEVFRKRKRKHLVFEYCDHTLLNELERNPNGVADGVKSVLWQTLOALNFCIH 120
DB 61 PNLVNLIEVFRKRKRKHLVFEYCDHTLLNELERNPNGVADGVKSVLWQTLOALNFCIH 120
QY 121 NCIHREDIKPENILITKQGIKICDFGFAQILIPGDATDYVATRWYRAPELLVGDTOYGS 180
DB 121 NCIHREDIKPENILITKQGIKICDFGFAQILIPGDATDYVATRWYRAPELLVGDTOYGS 180
QY 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLVETGFRHVDQAGLELLTSSDPPA 240
DB 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLVETGFRHVDQAGLELLTSSDPPA 240
QY 241 VASQSAGITGKLI PRHQSI FKSNGFFHGISIPEDMETLEEKFSVDHPVALNFMKGCJLK 300
DB 218 -----GKLI PRHQSI FKSNGFFHGISIPEDMETLEEKFSVDHPVALNFMKGCJLK 268
QY 301 MNPDRLTCSQLLESSYFDSFQEAQIKRKARNEGRRRRQ--QVLPL 345
DB 269 MNPDRLTCSQLLESSYFDSFQEAQIKRKARNEGRRRRRQCNQLPL 315

RESULT 5

AAE00492

ID AAE00492 standard; protein; 324 AA.

XX

AC AAE00492;

XX

DT 19-JUN-2001 (first entry)

DE Human kinase #3.

XX

KW Human; kinase; gene therapy; bioreactor; mental disorder;

KW biological disorder.

XX

OS Homo sapiens.

XX WC200123579-A1.
PN
XX
PD 05-APR-2001.
XX
PF 27-SEP-2000; 2000WC-US26621.
XX
PR 28-SEP-1999; 99US-0156511.
XX
PA (LEXI-1) LEXICON GENETICS INC.
XX
PI Donoh G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
XX
XX WPI; 2001-266166/27.
DR N-PSDB; AAC03814.
XX
XX New isolated human kinase polynucleotide useful for generating
PT antibodies, as reagents in diagnostic assays and for screening for
PT compounds useful for treating mental, biological or medical diseases -
XX
PS Disclosure; Page 30; 38pp; English.

CC The present sequence is novel human protein
CC (NHP) known as human kinase. The human kinases share structural
CC similarity with animal kinases, more particularly serine or
CC threonine protein kinases. Human kinase cDNA is useful for the
CC detection of mutant human kinase for the diagnosis of disease,
CC and also as a therapeutic. It is useful for screening drugs
CC effective in the treatment of symptomatic or phenotypic
CC manifestations perturbing the normal function of NHP in the
CC body. The NHP nucleotide sequences are useful for generation of
CC antibodies, as reagents in diagnostic assays, for the
CC identification of other cellular gene products related to human
CC kinases, and as reagents in assays for screening compounds that
CC are useful for treating mental, biological or medical disorders.
CC NHP oligonucleotides are used as probes. The labelled NHP probes
CC are useful for screening human genomic library for identifying
CC polymorphisms and as primers in amplification assays to detect
CC mutations within the exons, introns and splice sites that can
CC be used in diagnostics and pharmacogenomics. Nucleotide construct
CC encoding NHP products are used to genetically engineer cells
CC in vivo that functions as bioreactors in the body delivering a
CC continuous supply of NHP to the body. Nucleotide constructs
CC encoding functional NHPs are used in gene therapy for the
CC modulation of NHP expression.

XX SQ Sequence 324 AA;

Query Match 88.6%; Score 1612; DB 22; Length 324;
Best Local Similarity 90.6%; Pred. No. 3.4e-164;
Matches 310; Conservative 0; Mismatches 0; Indels 32; Gaps 1;

QY 1 MEKYEKLAKTGEISYGVVFKCRNKTSGQVAVKFKVESEDDPVVKKIALREIRMLKQLKH 60
DB 1 MEKYEKLAKTGEISYGVVFKCRNKTSGQVAVKFKVESEDDPVVKKIALREIRMLKQLKH 60
QY 61 PNLVNLIEVFRKRKRKHLVFEYCDHTLLNELERNPNGVADGVKSVLWQTLOALNFCIH 120
DB 61 PNLVNLIEVFRKRKRKHLVFEYCDHTLLNELERNPNGVADGVKSVLWQTLOALNFCIH 120
QY 121 NCIHREDIKPENILITKQGIKICDFGFAQILIPGDATDYVATRWYRAPELLVGDTOYGS 180
DB 121 NCIHREDIKPENILITKQGIKICDFGFAQILIPGDATDYVATRWYRAPELLVGDTOYGS 180
QY 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLVETGFRHVDQAGLELLTSSDPPA 240
DB 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLVETGFRHVDQAGLELLTSSDPPA 240
QY 241 VASQSAGITGKLI PRHQSI FKSNGFFHGISIPEDMETLEEKFSVDHPVALNFMKGCJLK 300
DB 218 -----GKLI PRHQSI FKSNGFFHGISIPEDMETLEEKFSVDHPVALNFMKGCJLK 268
QY 301 MNPDRLTCSQLLESSYFDSFQEAQIKRKARNEGRRRRQV 342

Db 269 MNPDRLTCSQLLESSYFDSFQEAQIKRKARNEGRRARQCV 310
|||||
RESULT 6
AAU03525
ID AAU03525 standard; Protein; 360 AA.
XX
AC AAU03525;
XX
DT 12-SEP-2001 (first entry)
XX
DE Human protein kinase #25.
XX
KW Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
KW metabolic disorder; immune related disease; neurological disorder;
KW neurodegenerative disorder; inflammatory disorder; infectious disease;
KW reproductive disorder.
XX
OS Homo sapiens.
XX
PN WC200138503-A2.
XX
PD 31-MAY-2001.
XX
PF 22-NOV-2000; 2000WO-US32085.
XX
PR 24-NOV-1999; 99US-0167482.
XX
PA (SUGE-) SUGEN INC.
XX
PI Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R,
PI Flanagan P, Clary D;
XX
DR WPI; 2001-343950/36.
DR N-PSDB; AAS06725.
XX
PT Nucleic acids encoding human kinase polypeptides, useful for preventing
PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and
PT neuronal-associated diseases, and microbial infections -
XX
PS Claim 7; Figure 2; 433pp; English.
XX
CC AAU03501-AAU03557 represent novel human protein kinases #1-57. The
CC novel protein kinases have been identified as members of the tyrosine
CC or serine/threonine kinase (PTK and STK) families. The polynucleotides
CC encoding protein kinases and the polypeptides may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate kinase expression. For example, they may be used to treat
CC cancers (especially cancers of haematopoietic origin), cardiovascular
CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
CC immune related diseases (e.g. rheumatoid arthritis), neurological
CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).
CC Additionally, polynucleotides encoding protein kinases may be
CC used for gene therapy and as DNA probes in diagnostic assays.
CC The protein kinase polypeptides may be used as antigens in the production
CC of antibodies against the protein kinases and in assays to identify
CC modulators of protein kinase expression and activity.
XX
SQ Sequence 360 AA;
Query Match 87.5%; Score 1593; DB 22; Length 360;
Best Local Similarity 85.2%; Pred. No. 4.4e-162;
Matches 311; Conservative 2; Mismatches 0; Indels 52; Gaps 3;
QY 1 MEKYEKLAKTGECSYGWVFKCRNKTSGQVAVKKFVESDDPVVKKIALREIRMLKQLKH 60
|||||
Db 1 MEKYEKLAKTGECSYGWVFKCRNKTSGQVAVKKFVESDDPVVKKIALREIRMLKQLKH 60
QY 61 PNLVNLIEVFRKRKQHLVFEYCDHTLLNELEPNPVGADGVKSVLWQTLQALNFCIH 120
|||||

Db 61 PNLVNLIEVFRKRKQHLVFEYCDHTLLNELEPNPVGADGVKSVLWQTLQALNFCIH 120
QY 121 NCIHRDIKPENILITKGIIKICDFGFAQLL-----IPGDAYTDYVA 162
|||||
Db 121 NCIHRDIKPENILITKGIIKICDFGFAQLLSNSHVGRIDLPLDLIDAFAPVGDAYDYVA 180
|||||
QY 163 TRWYRAPELLVGDTOYGSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLVETGF 222
|||||
Db 181 TRWYRAPELLVGDTOYGSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTL----- 235
QY 223 RHVDQAGLELLTSSDPPAVASQAGITGKLI PRHQSI FKSNGFPHGISIPEPEMETLEE 282
|||||
Db 236 -----GKLI PRHQSI FKSNGFPHGISIPEPEMETLEE 268
QY 283 KFSDVHPVALNFMKGCCKMNPDDRLLTCSQLLESSYFDSFQEAQIKRKARNEGRRRRQ-- 340
|||||
Db 269 KFSDVHPVALNFMKGCCKMNPDDRLLTCSQLLESSYFDSFQEAQIKRKARNEGRRRRQQN 328
QY 341 QVLPL 345
|||||
Db 329 QLLPL 333
RESULT 7
AAB65643
ID AAB65643 standard; Protein; 296 AA.
XX
AC AAB65643;
XX
DT 27-MAR-2001 (first entry)
XX
DE Novel protein kinase, SEQ ID NO: 170.
XX
KW Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
KW immunosuppressive; cardiant; renal; antiinflammatory; antiasthmatic;
KW dermatological; antidiabetic; antiinfertility; gene therapy; vaccine;
KW immune disorder; cardiovascular disease; neurodegenerative disease;
KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;
KW inflammatory pelvic disease; multiple sclerosis; psoriasis.
XX
OS Mus musculus.
XX
PN WC2000073459-A2.
XX
PD 07-DEC-2000.
XX
PF 26-MAY-2000; 2000WO-US14842.
XX
PR 28-MAY-1999; 99US-0136503.
XX (SUGE-) SUGEN INC.
PI Plowman GD, Martinez R, Whyte D, Sudarsanam S;
XX
DR WPI; 2001-032161/04.
DR N-PSDB; AAF44670.
XX
PT Nucleic acids encoding kinase polypeptides, useful for diagnosing and
PT treating immune-related diseases and disorders, cardiovascular disease,
PT neurodegenerative diseases and/or cancers -
XX
PS Claim 10; Fig 1; 310pp; English.
XX
CC The present sequence is a novel protein kinase. The novel protein kinases
CC and the nucleic acids that encode them may be used in the treatment and
CC diagnosis of diseases associated with inappropriate kinase expression
CC such as immune-related diseases and disorders, cardiovascular disease,
CC neurodegenerative diseases and/or cancers. The nucleic acids and
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays. The kinase polypeptides may be used as antigens in the production
CC of antibodies of kinase expression and activity. Anti-kinase antibodies
CC and kinase antagonists may also be used to down regulate kinase
CC expression and activity. Diseases related to kinase expression and

CC activity include rheumatoid arthritis, atherosclerosis, autoimmune
CC disorders, complications of organ transplantation, myocardial infarction,
CC immune disorders, cardiomyopathies, strokes, renal failure,
CC oxidative-stress related disorders, chronic inflammatory bowel disease,
CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,
CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and
CC reproductive disorders.

XX
SQ Sequence 296 AA;
Query Match 69.8%; Score 1270.5; DB 22; Length 296;
Best Local Similarity 80.5%; Pred. No. 1.5e-127;
Matches 243; Conservative 19; Mismatches 5; Indels 35; Gaps 3;

QY 46 KIALREIRMLKQLKHNLVNLIIEVFRKRKRMHLVFEYCDHTLLNELERNPNVGADGVTKS 105
DB 1 KIALREIRMLK-LKHPNLVNLIEVFRKRKRMHLVFEYCDHTLLNELERNPNVGSDGVTKS 59

QY 106 VLWQTLQALNPFCHIHNCIHRDIKPENILITKQGIKICDFGFAQLIIPGDAYTDYVATRW 165
DB 60 VLWQTLQALNPFCHKHNCIHRDVKPNILITKQGIKICDFGFAQLIIPGDAYTDYVATRW 119

QY 166 YRAPELLVGDYQYSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIRTLVETGFRHV 225
DB 120 YRAPELLVGDYQYSSVDWAVGCVFAELLTGQPLWPGKSDVDQLYLIRTL----- 171

QY 226 DQAGLELLTSSDPPAVASQASGITKLIPIRHQSIFKSNGFPHGISIPEPEDMETLEEKFS 285
DB 172 -----GKLIPIRHQSIFRSNCFFRGISIPEPEDMETLEEKFS 207

QY 286 DVHPVALNFMKGCLKXNDDRLTCSQLLESSYFDSFQEAQIKRKARNEGRRRQ--QVL 343
DB 208 NVQPVALSFMKGCLKXNDFERLTCAQLLDSAYESFOEDQMKRKARSEGRSRQQNQLL 267

QY 344 PL 345
DB 268 PL 269

RESULT 8
ABP62954
ID ABP62954 standard; Protein; 358 AA.
XX
AC ABP62954;
XX 14-OCT-2002 (first entry)
XX Human polypeptide SEQ ID NO 391.
XX Human; vulnery; dermatological; neuroprotective; nootropic; cancer;
KW antiparkinsonian; immunostimulant; cytostatic; immunosuppressive;
KW antidiabetic; antiallergic; gene therapy; wound healing; tissue repair;
KW burn; central nervous system disorder; Alzheimer's disease;
KW parkinson's disease; Huntington's disease; immune disorder;
KW autoimmune disorder; multiple sclerosis; diabetes; allergy.

XX
OS Homo sapiens.
XX
XX WO200218424-A2.
XX
XX 07-MAR-2002.
XX
XX 31-AUG-2001; 2001WO-US27093.
XX
XX 01-SEP-2000; 2000US-0654935.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Asundi V, Zhou P, Xue AJ, Ren F, Zhang J, Wang J;
PI Zhao QA, Wang D, Liu C, Drmanac RT, Wehrman T;
XX WPI; 2002-583321/62.
DR N-PSDB; ABQ93433.

XX New polynucleotide and polypeptides, useful for treatment and diagnosis
PT of Alzheimer's, Parkinson's, Huntington's, amyotrophic lateral
PT sclerosis, immune deficiencies, cancer, autoimmune disorders, multiple
PT sclerosis, diabetes and allergies -
XX
XX Claim 20; SEQ ID NO 391; 284pp + Sequence Listing; English.

XX The invention relates to an isolated polynucleotide (I) comprising one of
CC 245 sequences (ABQ93288-ABQ93532). Treating a condition comprising
CC administering to a mammalian subject a composition comprising the protein
CC (I) encoded by (I) (ABP62809-ABP63053) or an antibody (II) to (I).
CC (I), (II) and (III) are useful for diagnostic evaluation of disorders.
CC (I) is useful for gene therapy of diseases and (II) can be used for
CC therapeutic treatment. Diseases that may be treated include wound healing
CC and tissue repair, burns, central nervous system disorders (e.g.
CC Alzheimer's, Parkinson's, Huntington's and amyotrophic lateral
CC sclerosis), immune deficiencies, cancer, autoimmune disorders, multiple
CC sclerosis, diabetes and allergies.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 358 AA;
SQ

Query Match 63.9%; Score 1162.5; DB 23; Length 358;
Best Local Similarity 63.8%; Pred. No. 8.3e-116;
Matches 213; Conservative 42; Mismatches 46; Indels 33; Gaps 2;

QY 1 MEKYEKLAKTGEGSYGVVFKCRNKTSGQVAVKVFVESEDDPVKKIALREIRMLKQLKH 60
DB 2 MEKYEKIGTIGEGSYGVVFKCRNRDTGQIVAIKKFLESEDDPVIKKIALREIRMLKQLKH 61

QY 61 PNLVNLIEVFRKRKRMHLVFEYCDHTLLNELERNPNVGADGVKSVLWQTLQALNFCIH 120
DB 62 PNLVNLIEVFRKRRLHLVFEYCDHTVLHEDRYQRGVPEHLVKSITWQTLQAVNFCCHKH 121

QY 121 NCIHRIKPNILITKQGIKICDFGFAQLI- PGDAYTDYVATRWYRAPELLVGDYQY 179
DB 122 NCIHRIKPNILITKHSVIKLCDFGFAQLITGSPDYTDYVATRWYRSPPELLVGDYQY 181

QY 180 SSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIRTLVETGFRHVDQAGLELLTSSDPP 239
DB 182 PPVDVWAIGCVFAELLSGVPLWPGKSDVDQLYLIRKTL----- 219

QY 240 AVASQSAGITKLIPIRHQSIFKSNGFPHGISIPEPEDMETLEEKFSVHPVALNPMKGCL 299
DB 220 -----GDLIPRHQQVFSTNQYFSGVKIPDPEDMEPLEKFPNISYPALGLLKGCL 269

QY 300 KMNPDDRLLTCSQLLESSYFDSFQEAQIKRKARNE 333
DB 270 HMDPTERTLTCQLLHPYFENIREIEDLAKEHNK 303

RESULT 9
AAG78547
ID AAG78547 standard; Protein; 228 AA.
XX
AC AAG78547;
XX
DT 09-MAR-2002 (first entry)
XX
DE Human kinase 14257 amino acid sequence.
XX
KW protein kinase; enzyme; cytostatic; osteopathic; hepatotropic;
KW antidiabetic; neuroprotective; antiarthritic; dermatological;
KW immunosuppressive; antiinflammatory; antithyroid; antipsoriatic;
KW ophthalmological; antiallergic; antiasthmatic; antiatherosclerotic;
KW hypotensive; vasotropic; antiarrhythmic; virucide; anorectic;
KW metabolic; immunomodulator; analgesic; cellular proliferative disorder;
KW cancer; acute lymphoblastic leukaemia; Hodgkin's disease;
KW bone metabolism disorder; osteoporosis; immune system disorder;
KW inflammatory; diabetes mellitus; osteoarthritis; asthma;

KW cardiovascular disorder; hypertension; coronary artery disease;
XW endothelial cell disorder; psoriasis.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 4..218
FT /note= "eukaryotic protein kinase domain"
FT Modified-site 9..15
FT /note= "tyrosine kinase phosphorylation site"
FT Modified-site 23..26
FT /note= "N-glycosylation site"
FT Modified-site 27..32
FT /note= "N-myristoylation site"
FT Modified-site 38..41
FT /note= "casein kinase II phosphorylation site"
FT Modified-site 97..102
FT /note= "N-myristoylation site"
FT Modified-site 132..134
FT /note= "serine/threonine protein kinase active-site
FT signal site"
FT 161..218
FT /note= "kinase transferase protein
FT serine/threonine-protein ATP-binding II phosphorylation
FT casein alpha chain domain"
FT Modified-site 180..183
FT /note= "casein kinase II phosphorylation site"
FT Modified-site 188..193
FT /note= "N-myristoylation site"
FT Modified-site 204..211
FT /note= "tyrosine kinase phosphorylation site"
FT Modified-site 205..208
FT /note= "casein kinase II phosphorylation site"
XX
PN WO200179488-A2.
XX
PD 25-OCT-2001.
XX
PF 13-APR-2001; 2001WO-US12186.
XX
PR 13-APR-2000; 2000US-196910P.
XX (MILL-) MILLENNIUM PHARM INC.
PA
XX Kapeller-Libermann R;
PI
XX
DR WPI; 2002-034355/04.
DR N-PSDB; AAI64248.
XX
PT New 14257 polypeptides (protein kinases), useful as diagnostic targets
PT and therapeutic agents for controlling cellular proliferative and/or
PT differentiative disorder, bone disorders, immune disorders and
PT cardiovascular disorders -
XX
PS Claim 1c; Fig 1A; 98pp; English.
XX
CC The invention relates to an isolated 14257 polypeptide and nucleic
CC acid encoding it. The 14257 protein is a protein kinase that acts as a
CC modulating agent in regulating a variety of cellular processes,
CC including cell proliferation, differentiation, growth and division.
CC The activity of the protein of the invention may be described as;
CC cytostatic; osteopathic; hepatotropic; antidiabetic; neuroprotective;
CC antiarthritic; dermatological; immunosuppressive; antiinflammatory;
CC antithyroid; antipsoriatic; ophthalmological; antiallergic;
CC antiasthmatic; antiatherosclerotic; hypotensive; vasotropic;
CC antiarrhythmic; virucide; anorectic; metabolic; immunomodulator and
CC analgesic. The protein of the invention may act as a novel diagnostic
CC target or therapeutic agent controlling certain disorders, for example
CC kinase-associated or other 14257-associated disorders. These may include
CC cellular proliferative disorders such as cancers e.g. acute lymphoblastic
CC leukemia or Hodgkin's disease. Other disorders include bone metabolism
CC disorders such as osteoporosis, disorders of the immune system, e.g.
CC inflammatory , diabetes mellitus, osteoarthritis and asthma. Proteins of

CC the invention may also be of use as therapeutic agents in cardiovascular
CC disorders such as hypertension and coronary artery disease, and some
CC endothelial cell disorders, including psoriasis. The current
CC sequence represents a human kinase 14257 amino acid sequence.
XX
SQ Sequence 228 AA;
Query Match 62.3%; Score 1134; DB 23; Length 228;
Best Local Similarity 98.2%; Pred. No. 4.7e-113;
Matches 213; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 MEKYEKLAKTGEISYGVVFKCRNKTSGVVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
DB 1 MGKYEKLAKTGEISYGVVFKCRNKTSGVVAVKKFVESEDDPIVKKIALREIRMLKQLKH 60
QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELERNPNGVADGVKSVLWQTLQALNFCHIH 120
DB 61 PNLVNLIEVFRKRKMHVFEYCDHALNELERNPNGVADGVKSVLWQTLQALNFCHIH 120
QY 121 NCIHRIKPNILITKQGIKICDFGFAQILIPGDAYTDYVATRWYRAPELLVGDTOYGS 180
DB 121 NCIHRIKPNILITKQGIKICDFGFAQILIPGDAYTDYVATRWYRAPELLVGDTOYGS 180
QY 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTL 217
DB 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTL 217
RESULT 10
AAB65642
ID AAB65642 standard; Protein; 247 AA.
XX
AC AAB65642;
XX
DT 27-MAR-2001 (first entry)
XX
DF Novel protein kinase, SEQ ID NO: 169.
XX
KW Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
KW immunosuppressive; cardiac; renal; antiinflammatory; antiasthmatic;
KW dermatological; antidiabetic; antiinfertility; gene therapy; vaccine;
KW immune disorder; cardiovascular disease; neurodegenerative disease;
KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;
KW inflammatory pelvic disease; multiple sclerosis; psoriasis.
XX
OS Homo sapiens.
XX
PN WO200073469-A2.
XX
PD 07-DEC-2000.
XX
PF 26-MAY-2000; 2000WO-US14842.
XX
PR 28-MAY-1999; 99US-0136503.
XX
PA (SUGEN) SUGEN INC.
XX
PI Plowman GD, Martinez R, Whyte D, Sudersanam S;
XX
DR WPI; 2001-032161/04.
DR N-PSDB; AAF44669.
XX
PT Nucleic acids encoding kinase polypeptides, useful for diagnosing and
PT treating immune-related diseases and disorders, cardiovascular disease,
PT neurodegenerative diseases and/or cancers -
XX
PS Claim 10; Fig 1; 310pp; English.
XX
CC The present sequence is a novel protein kinase. The novel protein kinases
CC and the nucleic acids that encode them may be used in the treatment and
CC diagnosis of diseases associated with inappropriate kinase expression
CC such as immune-related diseases and disorders, cardiovascular disease,
CC neurodegenerative diseases and/or cancers. The nucleic acids and

complementary sequences may also be used as DNA probes in diagnostic assays. The kinase polypeptides may be used as antigens in the production of antibodies of kinase expression and activity. Anti-kinase antibodies and kinase antagonists may also be used to down regulate kinase expression and activity. Diseases related to kinase expression and activity include rheumatoid arthritis, atherosclerosis, autoimmune disorders, complications of organ transplantation, myocardial infarction, immune disorders, cardiomyopathies, strokes, renal failure, oxidative-stress related disorders, chronic inflammatory bowel disease, chronic inflammatory pelvic disease, multiple sclerosis, asthma, osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and reproductive disorders.

Sequence 247 AA;
Query Match 61.7%; Score 1123; DB 22; Length 247;
Best Local Similarity 86.1%; Pred. No. 8.1e-112;
Matches 217; Conservative 1; Mismatches 0; Indels 34; Gaps 2;
QY 96 NGVADGVKSVLWQTLQALNFCIHNCIHRDIKPNILITKQGIKICDFGFAQILIPGD 155
Db 1 NGVADGVKSVLWQTLQALNFCIHNCIHRDIKPNILITKQGIKICDFGFAQILIPGD 60
QY 156 AYTDYVATRWYRAPPELLVGDTOYGSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIR 215
Db 61 AYTDYVATRWYRAPPELLVGDTOYGSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIR 120
QY 216 TLVETGFRHVDQAGLELLTSSDPPAVASQSAGITGKLIERHQSIKSNQFFHGISIPEPE 275
Db 121 TL-----GKLIERHQSIKSNQFFHGISIPEPE 149
QY 276 DMETLEEKFSVDHPVALNFMKGCLKNPDDRLTCSQLLESSYFDSFQEAQIKRKARNEGR 335
Db 149 DMETLEEKFSVDHPVALNFMKGCLKNPDDRLTCSQLLESSYFDSFQEAQIKRKARNEGR 208
QY 336 NRRRQ--QVLPL 345
Db 209 NRRRQNQLLPL 220

RESULT 11
ABB63118
ID ABB63118 standard; Protein; 392 AA.
XX ABB63118;
AC ABB63118;
XX 26-MAR-2002 (first entry)
XX Drosophila melanogaster polypeptide SEQ ID NO 16146.
XX Drosophila; developmental biology; cell signalling; insecticide;
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX Drosophila melanogaster.
XX WO200171042-A2.
PN 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US09231.
PF 23-MAR-2000; 2000US-191637P.
XX 11-JUL-2000; 2000US-061415C.
XX (PEXE) PE CORP NY.
PA Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX N-PSDB; ABL0721.
XX New isolated nucleic acid detection reagent for detecting 1300 or more genes from Drosophila and for elucidating cell signalling and cell-cell

interactions -
XX Disclosure; SEQ ID NO 16146; 2lpp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072).
XX The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 392 AA;
Query Match 55.0%; Score 1001.5; DB 22; Length 392;
Best Local Similarity 53.0%; Pred. No. 1.9e-98;
Matches 184; Conservative 65; Mismatches 61; Indels 37; Gaps 3;
QY 1 MEKYEKLAKTGEESYGVWFKCRNKTSGOVVAVKKEVESEDDPVVKIALREIRMLKQLKH 60
Db 1 MORYEKL SRLGEGSYGVVYKCRDRETGALVAVKRFVESEDDPAIRKIALREIRLLKNLKH 60
QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELRNPNGVADGVKSVLWQTLQALNFCIH 120
Db 61 PNLVSLLEVFRRKRLHLVFEFCELTVLHELRHPQCGPEHLTKQICYQTLGVAYCHKQ 120
QY 121 NCIHARDIKPENILITKQGIKICDFGFAQILIPGDAYTDYVATRWYRAPPELLVGDTOYGS 180
Db 121 GCLHRDIKPENILITAGQGVKLCDFGFARMLSPGENYTDYVATRWYRAPPELLVGDTOYGT 180
QY 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLVETGFRHVDQAGLELLTSSDPPA 240
Db 181 PVDVWAIGCLFAELVRGEALWPGRSVDQLYLIRKTL----- 217
QY 241 VASQSAGITGKLIPRHQSIFKSNQFFHGISIPEPEDMETLEEKF--SDVHPVALNFMKG 297
Db 218 -----GDLPRHIQIFGQNEYFKGITLPVPTPLEPLEDKMPAKSQQNPLTIDFLKX 268
QY 298 CLKNPDDRLTCSQLLESSYFDSF--QEAQIKRKARNEGRNRRRQOV 342
Db 269 CLDKDPTKRWSCBKLTKHSYFDDYIAKQRELEHVNSLEAANLRQOOL 315

RESULT 12
ABU11689
ID ABU11689 standard; Protein; 197 AA.
XX ABU11689;
AC ABU11689;
XX 13-FEB-2003 (first entry)
XX Human MDDT polypeptide SEQ ID 636.
DE MDDT; human; disease detection and treatment molecule polypeptide;
XX anti-inflammatory; immunosuppressive; osteopathic; cytostatic; anti-HIV;
KW haemostatic; nephrotropic; antianaemic; antipsoriatic; hepatotropic;
KW gene therapy; protein replacement therapy; cell proliferative disorder;
KW cancer; adenocarcinoma; leukaemia; lymphoma; melanoma; myeloma; sarcoma;
KW anaemia; Crohn's disease; acquired immunodeficiency syndrome; AIDS;
KW Goodpasture's syndrome; inflammation; osteoporosis; thrombocytopaenia;
KW psoriasis; hepatitis.
XX Homo sapiens.
XX WO200279449-A2.
PN 10-OCT-2002.
XX 27-MAR-2002; 2002WO-US09944.

XX 28-MAR-2001; 2001US-279619P.
PR 29-MAR-2001; 2001US-280067P.
PR 29-MAR-2001; 2001US-280068P.
PR 16-MAY-2001; 2001US-291280P.
PR 17-MAY-2001; 2001US-291829P.
PR 17-MAY-2001; 2001US-291849P.
PR 19-JUN-2001; 2001US-299428P.
PR 20-JUN-2001; 2001US-299776P.
PR 20-JUN-2001; 2001US-300001P.
XX (INCY-) INCYTE GENOMICS INC.
XX Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
P: Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshey SR;
P: Daugherty SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstin EH;
P: Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
P: Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;
XX WPI; 2003-058431/05.
DR N-PSDB; ABX34679.
XX New purified disease detection and treatment molecule proteins and
PT polynucleotides, useful for diagnosing, treating or preventing cancers
P: (e.g. leukemia or sarcoma), anemia, Crohn's disease, AIDS, osteoporosis
PT or hepatitis -
XX Claim 27; SEQ ID NO 636; 339pp + Sequence Listing; English.
XX This invention describes a novel disease detection and treatment molecule
CC polypeptide (MDPT) which has anti-inflammatory, immunosuppressive,
CC osteopathic, cytostatic, anti-HIV, haemostatic, nephrotropic,
CC antianaemic, antipsoriatic and hepatotropic activity. The polynucleotides
CC and the polypeptides of the invention can be used for gene therapy,
CC protein replacement therapy and are useful for treating a variety of
CC diseases or conditions. These polypeptides or polynucleotides are
CC particularly useful for diagnosing, treating or preventing cell
CC proliferative disorders (e.g. cancers including adenocarcinoma,
CC leukaemia, lymphoma, melanoma, myeloma or sarcoma), anaemia, Crohn's
CC disease, acquired immunodeficiency syndrome (AIDS), Goodpasture's
CC syndromes, inflammation, osteoporosis, thrombocytopaenia, psoriasis or
CC hepatitis. ABU11450-ABU11845 represent the MDPT polynucleotides encoded
CC by ABU11450-ABU11845, described in the disclosure of the invention.
CC NOTE: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format from WIPC at
CC ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 197 AA;
SQ Query Match 52.1%; Score 948; DB 24; Length 197;
Best Local Similarity 98.9%; Pred. No. 3.7e-93;
Matches 180; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MEKYEKLAKTGECSYGWFKCRNKTSGQWVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
DB 6 MEKYEKLAKTGECSYGWFKCRNKTSGQWVAVKKFVESEDDPVVKKIALREIRMLKQLKH 65
QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEARNPVGADGVKSVLWQTLQALNFCIH 120
DB 66 PNLVNLIEVFRKRKMHVFEYCDHTLLNGLEARNPVGADGVKSVLWQTLQALNFCIH 125
QY 121 NCIHARDIKPENILITKQGIKICDFGFAQILIPGDAYTDYVATRWYRAPELLVGTQYGS 180
DB 126 NCIHARDIKPENILITKQGIKICDFGFAQILIPGDAYTDYVATRWYRAPELLVGTQYGS 185
QY 181 SV 182
DB 186 SV 187
RESULT 13
AAAY90724
ID AAY90724 standard; Protein: 566 AA.

XX AAY90724;
XX 15-AUG-2000 (first entry)
XX Rabbit KKIAMRE kinase SEQ ID NO:4.
DE Rabbit; KKIAMRE kinase; learning-induced kinase; learning; memory;
KW cdc2-related kinase; brain; gene therapy; genetic disorder; detection;
KW identification.
XX Oryctolagus cuniculus.
XX WO2000020567-A2.
PN 13-APR-2000.
PD 01-OCT-1999; 99WO-US23010.
XX 02-OCT-1998; 98US-0102906.
XX (UYSC-) UNIV SOUTHERN CALIFORNIA.
PA Thompson RF, Gomi H, Sun W;
P: WPI; 2000-328932/28.
XX N-PSDB; AAA29745.
DR Novel learning induced kinase polynucleotides and polypeptides, useful
PT for the analysis of learning and memory, and for gene therapy -
XX Claim 1; Fig 4; 64pp; English.
XX The present sequence represents a learning-induced kinase, designated
CC KKIAMRE kinase, which is isolated from rabbit brain tissue. KKIAMRE
CC kinase is a cdc2-related kinase. The KKIAMRE kinase polynucleotides can
CC be used to express recombinant protein for analysis, characterisation or
CC therapeutic use, as markers for tissues in which the protein is
CC preferentially expressed, as molecular weight markers on Southern gels,
CC as chromosome markers or tags, to compare endogenous DNA sequences in
CC patients to identify potential genetic disorders, as probes to hybridise
CC and discover novel related sequences, as a source of PCR primers, and as
CC an antigen to induce anti-DNA antibodies. The polypeptides can be used
CC in assay to discover biological activity, to raise antibodies, as tissue
CC markers, and to isolate correlative receptors or ligands. The
CC polynucleotides may also be used for gene therapy for the treatment of
CC disorders which are mediated by KKIAMRE kinase.
XX Sequence 566 AA;
SQ Query Match 51.7%; Score 941.5; DB 21; Length 566;
Best Local Similarity 49.7%; Pred. No. 9.7e-92;
Matches 177; Conservative 67; Mismatches 65; Indels 47; Gaps 6;
QY 1 MEKYEKLAKTGECSYGWFKCRNKTSGQWVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
DB 1 MEKYEKLAKTGECSYGWFKCRNKTSGQWVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEARNPVGADGVKSVLWQTLQALNFCIH 120
DB 62 ENLVNLLIEVCKKRWLVFEFVDHTLDDLELFPNGLDQVQVQYLFQINGIGFCHSH 120
QY 121 NCIHARDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGTQY 179
DB 121 NCIHARDIKPENILVSQSGWVKLDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGTQY 180
QY 180 SSVDIWAIGCVFAELITQPLWPGKSDVDQYLIIRTLVETGFRHVDQAGLELLTSSDPP 239
DB 181 KAVDVWAIGCLVTEMLMGEPLFPDSDIDQLYLMRCL----- 218
QY 240 AVASQSAGITGKLIPIRHSQIFKSNNGFFHGHSIPEPEMETLEEKFSVDVHPVAFNMKGCL 299
DB 219 -----GNLIPRQELFYKPNVFAGVRLPEIKSEPLRERYPKLSEWIDLAKKCL 268

CC NHP oligonucleotides are used as probes. The labelled NHP probes
CC are useful for screening human genomic library for identifying
CC polymorphisms and as primers in amplification assays to detect
CC mutations within the exons, introns and splice sites that can
CC be used in diagnostics and pharmacogenomics. Nucleotide construct
CC encoding NHP products are used to genetically engineer cells
CC in vivo that functions as bioreactors in the body delivering a
CC continuous supply of NHP to the body. Nucleotide constructs
CC encoding functional NHPs are used in gene therapy for the
CC modulation of NHP expression.

XX
SQ Sequence 187 AA;

Query Match 43.5%; Score 791; DB 22; Length 187;
Best Local Similarity 100.0%; Pred. No. 2.5e-76;
Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEKYEKLAKTGESGYGVFKCRNKTSGQWAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
DB |||||
QY 1 MEKYEKLAKTGESGYGVFKCRNKTSGQWAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
DB |||||
QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELERNPNGVADGVKSVLWQTLQALNFCIH 120
DB |||||
QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELERNPNGVADGVKSVLWQTLQALNFCIH 120
DB |||||
QY 121 NCIHRIKIPENILITKQGIKICDFGFAQIL 151
DB |||||
QY 121 NCIHRIKIPENILITKQGIKICDFGFAQIL 151
DB |||||

Search completed: November 13, 2003, 14:18:47
Job time : 41.8369 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 14:17:22 ; Search time 15.7251 Seconds
(without alignments)
933.660 Million cell updates/sec

Title: US-09-671-050-10
Perfect score: 1820
Sequence: 1 MEKYEKLAKTGECSYGVVFK.....EKARNEGRRRRQQVLPLKS 347

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: C
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/protdata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/protdata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/protdata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/protdata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/protdata/1/iaa/PC-US_COMB.pep:*
6: /cgn2_6/protdata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1157.5	63.6	358	US-09-411-628-11	Sequence 11, Appl
2	941.5	51.7	566	US-09-411-628-4	Sequence 4, Appli
3	915.5	50.3	493	US-09-411-628-10	Sequence 10, Appl
4	563	30.9	544	US-09-417-197-113	Sequence 113, App
5	557	30.6	298	US-08-874-347-25	Sequence 25, Appl
6	557	30.6	298	US-08-969-106-2	Sequence 2, Appli
7	557	30.6	298	US-09-093-522-25	Sequence 25, Appl
8	557	30.6	298	US-09-338-225-2	Sequence 2, Appli
9	557	30.6	298	US-09-266-225D-14	Sequence 14, Appl
10	557	30.6	544	US-09-417-197-115	Sequence 115, App
11	556	30.5	298	US-09-411-628-13	Sequence 13, Appl
12	554	30.4	297	US-08-176-620A-16	Sequence 16, Appl
13	554	30.4	297	US-08-461-985-16	Sequence 16, App
14	554	30.4	298	US-09-457-040B-29	Sequence 29, Appl
15	551.5	30.3	294	US-08-874-347-26	Sequence 26, Appl
16	551.5	30.3	290	US-09-093-522-26	Sequence 26, App
17	547.5	30.1	274	US-07-857-224B-31	Sequence 31, Appl
18	545	29.9	297	US-08-874-347-24	Sequence 24, Appl
19	545	29.9	297	US-09-093-522-24	Sequence 24, Appl
20	545	29.9	297	US-09-411-628-12	Sequence 12, App
21	541	29.7	297	US-08-874-347-23	Sequence 23, Appl
22	541	29.7	297	US-09-093-522-23	Sequence 23, App
23	540	29.7	274	US-08-318-947A-20	Sequence 20, Appl
24	540	29.7	274	US-08-795-303-20	Sequence 20, Appl
25	538.5	29.6	300	US-08-874-347-10	Sequence 10, Appl
26	538.5	29.6	300	US-09-093-522-10	Sequence 10, Appl
27	526	28.9	292	US-08-154-915-2	Sequence 2, Appli

28	526	28.9	292	2	US-08-464-517-38	Sequence 38, Appl
29	526	28.9	292	2	US-08-246-361A-38	Sequence 38, Appl
30	526	28.9	292	3	US-08-463-772-38	Sequence 38, Appl
31	526	28.9	292	5	PCT-US93-09945-2	Sequence 2, Appli
32	526	28.9	297	3	US-08-932-787B-21	Sequence 21, Appl
33	526	28.9	297	3	US-08-932-012C-21	Sequence 21, Appl
34	526	28.9	297	4	US-08-888-818C-21	Sequence 21, Appl
35	498.5	27.4	297	2	US-08-874-347-22	Sequence 22, Appl
36	498.5	27.4	297	3	US-09-093-522-22	Sequence 22, Appl
37	493.5	27.1	274	2	US-07-857-224B-30	Sequence 30, App-
38	491.5	27.0	527	4	US-09-739-455-2	Sequence 2, Appli
39	489.5	26.9	282	1	US-08-318-947A-19	Sequence 19, Appl
40	489.5	26.9	282	2	US-08-795-303-19	Sequence 19, Appl
41	489	26.9	323	2	US-08-874-347-21	Sequence 21, Appl
42	489	26.9	323	3	US-09-093-522-21	Sequence 21, Appl
43	488.5	26.8	317	1	US-08-463-090B-9	Sequence 9, Appli
44	488.5	26.8	317	2	US-08-874-347-18	Sequence 18, Appl
45	488.5	26.8	317	3	US-09-093-522-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1
US-09-411-628-11
; Sequence 11, Application US/09411628
; Patent No. 6428994
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; TITLE OF INVENTION: CDNA, GENOMIC, AND PREDICTED PROTEIN
; FILE OF INVENTION: SEQUENCES OF LEARNING-INDUCED KINASES
; FILE REFERENCE: 13761-707
; CURRENT APPLICATION NUMBER: US/09/411,628
; CURRENT FILING DATE: 1999-10-01
; EARLIER APPLICATION NUMBER: US 60/102,906
; EARLIER FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-411-628-11

Query Match	63.6%;	Score	1157.5;	DB	4;	Length	358;		
Best Local Similarity	64.7%;	Pred. No.	9.5e-114;						
Matches	211;	Conservative	41;	Mismatches	41;	Indels	33;	Gaps	2;
QY	1	MEKYEKLAKTGECSYGVVFKCRNKTSGQVAVKXPFVESEDDPVVKKIALREIRMLKQLKH	60						
Db	2	MEKYEKIGKIGEGSYGVVFKCRNRDTGQIVAIKKFLESEDDPVIKKIALREIRMLKQLKH	61						
QY	61	PNLVNLIIEVFRRRKMHVFEYCDHTLLNELERNPNGVADGVIKSVLWOTLOALNFCIH	120						
Db	62	PNLVNLLIEVFRRRKRLHLVFEYCDHTVLHELDYRGQVPEHLVKSITWOTLOAVNFCCHK	121						
QY	121	NCIHRDIKPENILITKQGIKICDFGFAQILI - PGDAYTDYVATRWYRAPELLVGDTQYG	179						
Db	122	NCIHRDVKPENILITKHSVIKLCDFGFARLLTGPSDYTYTDYVATRWYRSPELLVGDTQYG	181						
QY	180	SSVDIWAIGCVFAELLTGQPLWPGKSDVDQYLIIRTLVETGFRHVDQAGLELLTSSDPP	239						
Db	182	PPVDVWAIGCVFAELLSGVPLWPGKSDVDQYLIIRKTL-----	219						
QY	240	AVASQSAGITGKLIPIRHQSIPKSNGFFHGISIPEPEDMETLEEKFSVDHPVALNFMKGCL	299						
Db	220	-----GDLIPRHQQVFSTNQYFSGVKIPDPEDMEPLEKFPNISYPALGLLKGL	269						
QY	300	KMPDDRLTCSQLLESSYFDSFQEAQ	325						
Db	270	HMDPTERTCEQLLHHPYFENIREIE	295						

RESULT 2
US-09-411-628-4
; Sequence 4, Application US/09411628
; Patent No. 6428994
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; TITLE OF INVENTION: CDNA, GENOMIC, AND PREDICTED PROTEIN
; TITLE OF INVENTION: SEQUENCES OF LEARNING-INDUCED KINASES
; FILE REFERENCE: 13761-707
; CURRENT APPLICATION NUMBER: US/09/411,628
; CURRENT FILING DATE: 1999-10-01
; EARLIER APPLICATION NUMBER: US 60/102,906
; EARLIER FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 566
; TYPE: PR1
; ORGANISM: Oryctolagus cuniculus
US-C9-411-628-4

Query Match: 51.7%; Score 941.5; DB 4; Length 566;
Best Local Similarity 49.7%; Pred. No. 1.2e-90;
Matches 177; Conservative 67; Mismatches 65; Indels 47; Gaps 6;
QY 1 MEKYEKLAKTGESYGVVFKCRNKTSGQVAVKKFVESEDDPVYKIALREIRMLKQLKH 60
DB 1 MEKYENLGLVGEYSYGMVKCRNKDSGRIVAIAKKFLESDDDQVVKKIAMREIKLKLQLRH 60
QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEARNPNGVADGVKSVLWQTLQALNFCCHI 120
DB 61 ENLVNLEVCCKKRWYLVFEFVDHTILDLELFPNGLDQVVKYLFQIINGIGFCHSH 120
QY 121 NC1HRDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDTCYQ 179
DB 121 NI1HRDIKPENILVSQGVVKLCDFGFARTLAAPGEVYTDYVATRWYRAPELLVGDVKY 180
QY 190 SSVDIWAIGCVFAELLTGQPLWPGKSDVDQYLIIRTLVETGFRHVDQAGLELLTSSDPP 239
DB 181 KAVDVWAIGCLVTEMLMGEPLFPFGSDIDQYLIIRCL----- 218
QY 240 AVASQSAGITGKLIPRHQSIFKSNFGFHGISPEPEDMETLEEFSDVHPVALNFMKGCL 299
DB 219 -----GNLIPRHQELFYKNPVFAGVRLPEIKEREPLERRYPKLSEVVIDLAKKCL 268
QY 300 KMNPDRLTCSQLLESSYF--DSF-----QEAQIK--RKARN-----EGRNRRRQ 341
DB 269 HVDPDKRPFCALLHDFQMDGFAERFSQELQKVKQKDARNVSLSKSQNRKKEK 324

RESULT 3
US-09-411-628-10
; Sequence 10, Application US/09411628
; Patent No. 6428994
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; TITLE OF INVENTION: CDNA, GENOMIC, AND PREDICTED PROTEIN
; TITLE OF INVENTION: SEQUENCES OF LEARNING-INDUCED KINASES
; FILE REFERENCE: 13761-707
; CURRENT APPLICATION NUMBER: US/09/411,628
; CURRENT FILING DATE: 1999-10-01
; EARLIER APPLICATION NUMBER: US 60/102,906
; EARLIER FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 493
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-411-628-10

Query Match: 50.3%; Score 915.5; DB 4; Length 493;
Best Local Similarity 48.3%; Pred. No. 5.5e-88;

Matches 172; Conservative 68; Mismatches 69; Indels 47; Gaps 6;
QY 1 MEKYEKLAKTGESYGVVFKCRNKTSGQVAVKKFVESEDDPVYKIALREIRMLKQLKH 60
DB 1 MEKYENLGLVGEYSYGMVKCRNKDTGRIVAIAKKFLESDDDQVVKKIAMREIKLKLQLRH 60
QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEARNPNGVADGVKSVLWQTLQALNFCCHI 120
DB 61 ENLVNLEVCCKKRWYLVFEFVDHTILDLELFPNGLDYQVVKYLFQIINGIGFCHSH 120
QY 121 NC1HRDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDTCYQ 179
DB 121 NI1HRDIKPENILVSQGVVKLCDFGFARTLAAPGEVYTDYVATRWYRAPELLVGDVKY 180
QY 180 SSVDIWAIGCVFAELLTGQPLWPGKSDVDQYLIIRTLVETGFRHVDQAGLELLTSSDPP 239
DB 181 KAVDVWAIGCLVTEMLMGEPLFPFGSDIDQYLIIRCL----- 218
QY 240 AVASQSAGITGKLIPRHQSIFKSNFGFHGISPEPEDMETLEEFSDVHPVALNFMKGCL 299
DB 219 -----GNLIPRHQELFNKNPVFAGVRLPEIKEREPLERRYPKLSEVVIDLAKKCL 268
QY 300 KMNPDRLTCSQLLESSYF--DSF-----QEAQIK--RKARN-----EGRNRRRQ 341
DB 269 HVDPDKRPFCALLHDFQMDGFAERFSQELQKVKQKDARNVSLSKSQNRKKEK 324

RESULT 4
US-09-417-197-113
; Sequence 113, Application US/09417197
; Patent No. 6518021
; GENERAL INFORMATION:
; APPLICANT: Ole Thastrup, et al.
; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An
; TITLE OF INVENTION: On A Cellular Response
; FILE REFERENCE: 3759-0110P
; CURRENT APPLICATION NUMBER: US/09/417,197
; CURRENT FILING DATE: 1999-10-07
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 113
; LENGTH: 544
; TYPE: PR1
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: CDK2-EGFP fusion
US-09-417-197-113
Query Match: 30.9%; Score 563; DB 4; Length 544;
Best Local Similarity 37.2%; Pred. No. 1.1e-50;
Matches 124; Conservative 58; Mismatches 95; Indels 56; Gaps 8;
QY 1 MEKYEKLAKTGESYGVVFKCRNKTSGQVAVKKFVESEDDPVYKIALREIRMLKQLKH 60
DB 1 MENFQKVEKIGEGTYGVVYKARNKLTGEVVALKKIRLDTETEGVPSTAIRESISLKLNLH 60
QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEARNP-NGVADGVKSVLWQTLQALNFCCHI 119
DB 61 PNLVKLLDVHTENKLYJVFELHQDLKKFMDASALTGIPPLIKSYLFQLQLGLAFCHS 120
QY 120 HNCIHRDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDTCY 178
DB 121 HRVJHRDLKPKQNLINTEGAIKLADFGLARAFGVPVRYTHEVVTWYRAPELLGSKYY 180
QY 179 GSSVDIWAIGCVFAELLTGQPLWPGKSDVDQYLIIRTL----- 217
DB 181 STAVDIWSLGCIFAEMVTRRALFPDSEIDQLFRIFRTLGTPEVWPGVTSMVDYKPSF 240
QY 218 ---VETGFRHV---DQAGLELLTSS---DPPAVASQSAGITGKLIPRHQSIFKSNFGFH 267
DB 241 PKWARQDFSKVVPPLDEDGRSLLSQMLHYDPNKRISAKAAL-----AHPFFQ 287
QY 268 GISIPEPE-----DMETL-----BEKFSVDVHPV 290

Db 288 DVTKPVPHLRLWDPPVATMWSKGEELFTGVVP: 320

RESULT 5
US-08-874-347-25
; Sequence 25, Application US/08874347
; Patent No. 5863741
; GENERAL INFORMATION:
; APPLICANT: Limper, Andrew H.
; APPLICANT: Leof, Edward B.
; APPLICANT: Thomas, Charles F.
; APPLICANT: Gustafson, Michael P.
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
; TITLE OF INVENTION: CARINI
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C., P.A.
; STREET: 60 South Sixth Street, Suite 3300
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/874,347
; FILING DATE: 13-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ellinger, Mark S.
; REGISTRATION NUMBER: 34,812
; REFERENCE/DOCKET NUMBER: 07039/055001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-335-5070
; TELEFAX: 612-288-9696
; TELEX:
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 298 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-874-347-25

Query Match 30.6%; Score 557; DB 2; Length 298;
Best Local Similarity 38.4%; Pred. No. 1.8e-50;
Matches 118; Conservative 54; Mismatches 89; Indels 46; Gaps 6;
QY 1 MEKYEKLAKTGEYSYGVVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MENFQKVEKIGEGTYGVVYKARNKLTGEVVALKKIRLDTETEGVPSTAIRISLLKELNH 60
QY 61 PNLVNLIEVFRKRKXCHLVFEYCHTLLNELEARNP-NGVADGVKSVLWQTLQALNFCHI 119
Db 61 PNIIVKLLDVIHTENKLYLVFEFLHQDLKKFMDASALTGIPPLIKSYLFQLLQGLAFCHS 120
QY 120 HNCIHRDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDTQY 178
Db 121 HRVLRDLKPQNLLINTEGAIKLADFGLARAFGVPVRYTYTHEVTLWYRAPELLGSKYY 180
QY 179 GSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTL----- 217
Db 181 STAVDIWSLGCIFAEMVTRRALFPGDSEIDQLFRIFRTLGTpDEVVWPGVTSMPDYKPSF 240
QY 218 ---VETGFRHV---DQAGLELLTSS---DPPAVASQSAGITGKLI PRHQSIKSNNGFFH 267
Db 241 PKWARQDFSKVPPLEDGRSLLSQMLHYDPNKRISAKAAL-----AHPPFQ 287

QY 268 GISIPEP 274
Db 288 DVTKPVP 294
RESULT 6
US-09-969-106-2
; Sequence 2, Application US/08969106
; Patent No. 5986055
; GENERAL INFORMATION:
; APPLICANT: Yang, M.
; APPLICANT: Nandabalan, K.
; APPLICANT: Schulz, V.
; TITLE OF INVENTION: CDK2 INTERACTIONS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/969,106
; FILING DATE: 13-NOV-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7934-057
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 298 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-969-106-2

Query Match 30.6%; Score 557; DB 2; Length 298;
Best Local Similarity 38.4%; Pred. No. 1.8e-50;
Matches 118; Conservative 54; Mismatches 89; Indels 46; Gaps 6;
QY 2 MEKYEKLAKTGEYSYGVVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MENFQKVEKIGEGTYGVVYKARNKLTGEVVALKKIRLDTETEGVPSTAIRISLLKELNH 60
QY 61 PNLVNLIEVFRKRKXCHLVFEYCHTLLNELEARNP-NGVADGVKSVLWQTLQALNFCHI 119
Db 61 PNIIVKLLDVIHTENKLYLVFEFLHQDLKKFMDASALTGIPPLIKSYLFQLLQGLAFCHS 120
QY 120 HNCIHRDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDTQY 178
Db 121 HRVLRDLKPQNLLINTEGAIKLADFGLARAFGVPVRYTYTHEVTLWYRAPELLGSKYY 180
QY 179 GSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTL----- 217
Db 181 STAVDIWSLGCIFAEMVTRRALFPGDSEIDQLFRIFRTLGTpDEVVWPGVTSMPDYKPSF 240
QY 218 ---VETGFRHV---DQAGLELLTSS---DPPAVASQSAGITGKLI PRHQSIKSNNGFFH 267
Db 241 PKWARQDFSKVPPLEDGRSLLSQMLHYDPNKRISAKAAL-----AHPPFQ 287
QY 268 GISIPEP 274

QY 218 ---VETGFRHV---DQAGLELLTSS---DPPAVASQSAGITGKLI PRHQSIKSNNGFFH 267
DQ 241 PKWARQDFSKVWPPLDEGRSLSSQMLHYDPNKRISAKAAL-----AHPPFQ 287
QY 268 GISIPEP 274
Db 288 DVTKPVP 294

RESULT 9
US-09-266-225D-14
Query Match 30.6%; Score 557; DB 4; Length 298;
Best Local Similarity 38.4%; Pred. No. 1.8e-50;
Matches 118; Conservative 54; Mismatches 89; Indels 46; Gaps 6;
GENERAL INFORMATION:
APPLICANT: Nandabalan, Krishan
APPLICANT: Kingsmore, Stephen
APPLICANT: Tchernev, Velizar
TITLE OF INVENTION: Isolation and Characterization of Hermansky-Pudlak
TITLE OF INVENTION: Syndrome (HPS) Protein Complexes and HPS Protein-
TITLE OF INVENTION: Interacting Proteins
FILE REFERENCE: 15966-523
CURRENT APPLICATION NUMBER: US/09/266,225D
CURRENT FILING DATE: 1999-03-10
NUMBER OF SEQ ID NOS: 19
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 14
LENGTH: 298
TYPE: PRT
ORGANISM: Homo sapiens
US-09-266-225D-14

QY 1 MEKYEKLAKTGEISYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
DQ 1 MENFQKVEKIGEGTYGVVYKARNKLTGEVVALKKIRLDTETEGVPSTAIRISLLKELNH 60
QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEARNP-NGVADGVVKSVLWQTLQALNFCHI 119
DQ 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEARNP-NGVADGVVKSVLWQTLQALNFCHI 119
QY 120 HNCIHRDIKPENILITKQGIKICDFGPAQIL-IPGDAYTDVYVATRWYRAPELLVGDQY 178
DQ 121 HRVLRDLKPCNLLINTEGAIKLADFGCLARAFGVVPTVTHEVTLWYRAPEILLGSKYY 180
QY 179 GSSVDIWAIGCVFAELLTGQPLWPGKSDVDQYLIIRL----- 217
DQ 181 STAVDIWSLGCIFAEMVTRRALFPGDSEIDQLFRIFRITGTPDEVVWPGVTSYDYKPSF 240
QY 218 ---VETGFRHV---DQAGLELLTSS---DPPAVASQSAGITGKLI PRHQSIKSNNGFFH 267
DQ 241 PKWARQDFSKVWPPLDEGRSLSSQMLHYDPNKRISAKAAL-----AHPPFQ 287
QY 268 GISIPEP 274
Db 288 DVTKPVP 294

RESULT 10
US-09-417-197-115
Query Match 30.5%; Score 556; DB 4; Length 298;
Best Local Similarity 38.4%; Pred. No. 2.4e-50;
Matches 118; Conservative 54; Mismatches 89; Indels 46; Gaps 6;
GENERAL INFORMATION:
APPLICANT: Ole THASTRUP, et al.
TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An
TITLE OF INVENTION: On A Cellular Response
FILE REFERENCE: 3759-0110P
CURRENT APPLICATION NUMBER: US/09/417,197
CURRENT FILING DATE: 1999-10-07
NUMBER OF SEQ ID NOS: 143

SOFTWARE: Patentin version 3.0
SEQ ID NO 115
LENGTH: 544
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: EGFP-CDK2 fusion
US-09-417-197-115
Query Match 30.6%; Score 557; DB 4; Length 544;
Best Local Similarity 38.4%; Pred. No. 4.6e-50;
Matches 118; Conservative 54; Mismatches 89; Indels 46; Gaps 6;
QY 1 MEKYEKLAKTGEISYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
DQ 247 MENFQKVEKIGEGTYGVVYKARNKLTGEVVALKKIRLDTETEGVPSTAIRISLLKELNH 306
QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEARNP-NGVADGVVKSVLWQTLQALNFCHI 119
DQ 307 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEARNP-NGVADGVVKSVLWQTLQALNFCHI 119
QY 120 HNCIHRDIKPENILITKQGIKICDFGPAQIL-IPGDAYTDVYVATRWYRAPELLVGDQY 178
DQ 367 HRVLRDLKPCNLLINTEGAIKLADFGCLARAFGVVPTVTHEVTLWYRAPEILLGSKYY 426
QY 179 GSSVDIWAIGCVFAELLTGQPLWPGKSDVDQYLIIRL----- 217
DQ 427 STAVDIWSLGCIFAEMVTRRALFPGDSEIDQLFRIFRITGTPDEVVWPGVTSYDYKPSF 486
QY 218 ---VETGFRHV---DQAGLELLTSS---DPPAVASQSAGITGKLI PRHQSIKSNNGFFH 267
DQ 487 PKWARQDFSKVWPPLDEGRSLSSQMLHYDPNKRISAKAAL-----AHPPFQ 533
QY 268 GISIPEP 274
DQ 534 DVTKPVP 540
RESULT 11
US-09-411-628-13
Query Match 30.5%; Score 556; DB 4; Length 298;
Best Local Similarity 38.4%; Pred. No. 2.4e-50;
Matches 118; Conservative 54; Mismatches 89; Indels 46; Gaps 6;
GENERAL INFORMATION:
APPLICANT: University of Southern California
TITLE OF INVENTION: CDNA, GENOMIC, AND PREDICTED PROTEIN
TITLE OF INVENTION: SEQUENCES OF LEARNING-INDUCED KINASES
FILE REFERENCE: 13761-707
CURRENT APPLICATION NUMBER: US/09/411,628
CURRENT FILING DATE: 1999-10-01
EARLIER APPLICATION NUMBER: US 60/102,906
EARLIER FILING DATE: 1998-10-02
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13
LENGTH: 298
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(298)
OTHER INFORMATION: Xaa = Any Amino Acid
US-09-411-628-13
Query Match 30.5%; Score 556; DB 4; Length 298;
Best Local Similarity 38.4%; Pred. No. 2.4e-50;
Matches 118; Conservative 54; Mismatches 89; Indels 46; Gaps 6;
QY 1 MEKYEKLAKTGEISYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
DQ 1 MENFQKVEKIGEGTYGVVYKARNKLTGEVVALKKIRLDTETEGVPSTAIRISLLKELNH 60
QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEARNP-NGVADGVVKSVLWQTLQALNFCHI 119

Db 61 PNIIVKLLDVIHTENKLYLVFEFLHQGLKXFDASALTGIPLPLIKSYLFQLLQGLAFCHS 120
QY 120 HNCIHRDIKIPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDTOY 178
Db 121 HRVLHRLKPKQNLINTEGAIKLADFGGLARAFGVFVRTYTHEVVTLMWYRAPEILLGCKYV 180
QY 179 GSSVVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTL----- 217
Db 181 STAVDIWSLGCIFAEVMTTRALFPGDSEIDQLFRIFRTLGTDEVVWPGVTSMPOYKPSF 240
QY 218 ---VETGFRHV----DQAGLELLTSS---DPPAVASQSAGITGKLIPIRHQSIFKSNGFHF 267
Db 241 PKWARCDFSKVVPPLDEDGRSLLSQMLHYDPNKRISAKAAJ-----AHPEFQ 287
QY 268 GISIPEP 274
Db 298 DVTKPVP 294

RESULT 12
US-08-176-620A-16
; Sequence 16, Application US/08176620A
; Patent No. 5595904
; GENERAL INFORMATION:
; APPLICANT: Boulton, Teri G.
; APPLICANT: Cobb, Melanie H.
; APPLICANT: Yancopoulos, George D.
; APPLICANT: Nye, Steven
; APPLICANT: Panayotatos, Nikos
; TITLE OF INVENTION: A Family of Map2 Protein Kinases
; NUMBER OF SEQUENCES: 2;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/176,620A
; FILING DATE: 03-JAN-1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 6526-123
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 297 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-176-620A-16

Query Match 30.4%; Score 554; DB 1; Length 297;
Best Local Similarity 40.5%; Pred. No. 3.8e-50;
Matches 118; Conservative 58; Mismatches 73; Indels 42; Gaps 7;
QY 1 MEKYEKLAKTGECSYGVVFKCRNKTSGQVAVKKF-VESEDDPVVKKIALREIRMLKQLK 59
Db 1 MEDYTKIEKIGEGTYGVVYKGRHKTTGQVAVAMKKIRLESEEG-VPSTAIRESLJKEUR 59
QY 60 HPNLVNLIEVFRKRKMHVFEYCDHTLLNELEARNPG--VADGVIKSVLWQTLQALNFC 117

Db 60 HPNIVSLQDVLMDQSRLYLIFEFSMDLKKYLDISPPQGYMDSSLVKSYLYQILQIVFC 119
QY 118 HIHNCIHRDIKIPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGD 176
Db 120 HSRRLHRLKPKQNLIDDKGTIKLADFGGLARAFGPIRVVYTHYVVTLMWYRSPEVLGSA 179
QY 177 QYSSVVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTL----- 217
Db 180 RYSTPVDIWSIGTIFAELATKKPLFHGDEIDQLFRIFRALGTPNNEVWPEVESLQDYKN 239
QY 218 -----VETGFRHVDQAGLELLTSS---DPPAVASQSAGITGKLIPIRH 256
Db 240 TFPKWKPGSLASHVKNLDENGLDLSKMLIYDP-----AKRISGKMALNH 284

RESULT 13
US-08-461-985-16
; Sequence 16, Application US/08461985
; Patent No. 5872006
; GENERAL INFORMATION:
; APPLICANT: Boulton, Teri G.
; APPLICANT: Cobb, Melanie H.
; APPLICANT: Yancopoulos, George D.
; APPLICANT: Nye, Steven
; APPLICANT: Panayotatos, Nikos
; TITLE OF INVENTION: A Family of Map2 Protein Kinases
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,985
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/176,620
; FILING DATE: 03-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 6526-123
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 297 amino acids
; TYPE: aminc acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-461-985-16

Query Match 30.4%; Score 554; DB 2; Length 297;
Best Local Similarity 40.5%; Pred. No. 3.8e-50;
Matches 118; Conservative 58; Mismatches 73; Indels 42; Gaps 7;
QY 1 MEKYEKLAKTGECSYGVVFKCRNKTSGQVAVKKF-VESEDDPVVKKIALREIRMLKQLK 59
Db 1 MEDYTKIEKIGEGTYGVVYKGRHKTTGQVAVAMKKIRLESEEG-VPSTAIRESLJKEUR 59
QY 60 HPNLVNLIEVFRKRKMHVFEYCDHTLLNELEARNPG--VADGVIKSVLWQTLQALNFC 117

Db 60 HPNIVSLQDVLMQDSRLYLIFELFLSMDLKKYLDSPGQYMDSSLVKSXYLQILQGIWFC 119
Qy 118 HHCNHRDIKPEENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGT 176
Db 120 HSRVLRDLKPKQNLIDDKTIKADFLARAFGIPRVYTHYVTVLWYRSPVLLGSA 179
Qy 177 QYSSVDIWAIGCVFAELLTQPLWPGKSDVDQLYLIRTJ----- 217
Db 180 RYSTPVDIWSIGTIFAELATKKPLFHGDSEIDQLFRIFRALGTPNNEVWPEVESJQDYKN 239
Qy 218 -----VETGFRHVDQAGLELLTSS---DPPAVASQSAGITGKLIPRH 256
Db 240 TFPKWKPGSLASHVKNLDEGLDLSKMLIYDP-----AKRISGKMALNH 284

RESULT 14
US-09-457-040B-29
; Sequence 29, Application US/09457040B
; Patent No. 6387641
; GENERAL INFORMATION:
; APPLICANT: Vertex Pharmaceuticals Incorporated
; APPLICANT: Bellon, Steve
; TITLE OF INVENTION: Crystallized P38 Complexes
; FILE REFERENCE: VPI/98-14
; CURRENT APPLICATION NUMBER: US/09/457,040B
; CURRENT FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 29
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Human
US-09-457-040B-29

Query Match 30.4%; Score 554; DB 4; Length 298;
Best Local Similarity 38.4%; Pred. No. 3.8e-50;
Matches 118; Conservative 54; Mismatches 89; Indels 46; Gaps 6;
Qy 1 MEKYEKLAKTGECSYGWVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MENQKVEKIGEGTYGVVYKARNKLTGEVVALKKIRLDTETEGVPSTAIRISLKELNH 60
Qy 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELRNP-NGVADGVKSVLWQTLQALNFCIH 119
Db 61 PNIVKLLDVIHTENKLYLVFEFLHQDLKKFMDASALTGIPLEPLIKSYLFQLLQGLAFCHS 120
Qy 120 HNCIHRDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGTQY 178
Db 121 HRVLRDLKPKQNLINTEGAIKADFLARAFGVPVRYTTHEVWTLWYRAPELLGCKXY 180
Qy 179 GSSVDIWAIGCVFAELLTQPLWPGKSDVDQLYLIRTJ----- 217
Db 181 STAVDIWSLGCIFAEMVTRRALFPDSEIDQLFRIFRTLGTDEVVWPGVTSNPDYKPSF 240
Qy 218 ---VETGFRHV---DQAGLELLTSS---DPPAVASQSAGITGKLIPRHQSIFKSNGFH 267
Db 241 PKWARQDFSKVPPPLDEGDRSLLSQMLHYDPNKRISAKAAL-----AHFPFQ 267
Qy 268 GISIPEP 274
Db 288 DVTKVPV 294

RESULT 15
US-08-874-347-26
; Sequence 26, Application US/08874347
; Patent No. 5863741
; GENERAL INFORMATION:
; APPLICANT: Limper, Andrew H.
; APPLICANT: Leof, Edward B.
; APPLICANT: Thomas, Charles F.
; APPLICANT: Gustafson, Michael P.

; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
; TITLE OF INVENTION: CARINII
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C., P.A.
; STREET: 60 South Sixth Street, Suite 3300
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/874,347
; FILING DATE: 13-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ellinger, Mark S.
; REGISTRATION NUMBER: 34,812
; REFERENCE/DOCKET NUMBER: 07039/055001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-335-5070
; TELEFAX: 612-288-9696
; TELEX:
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 294 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-874-347-26

Query Match 30.3%; Score 551.5; DB 2; Length 294;
Best Local Similarity 34.8%; Pred. No. 6.9e-50;
Matches 116; Conservative 68; Mismatches 102; Indels 47; Gaps 8;
Qy 1 MEKYEKLAKTGECSYGWVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MEQYEKEKIGEGTYGVVYRARCVTNETIALKKIRLEQDEGVPSTAIRISLKEMHH 60
Qy 61 PNLVNLIEVFRKRKMHVFEYCDHTLLN-----ELERNPNGVADGVKSVLWQTLQAL 114
Db 61 GNVRLHDVHSEKRIYLVFEYLDLCLKKFMDSCEFAKNPT-----LIKSYLYQILRGV 115
Qy 115 NFCHIHNCIHRDIKPENILITKQ-GIHKICDFGFAQIL-IPGDAYTDYVATRWYRAPELL 172
Db 116 AYCHSHRVLHRDLKPCQNLIDRRTNALKLADFLARAFGIPVRTTTFHEVVTLWYRAPEIL 175
Qy 173 VGDQYGSVDIWAIGCVFAELLTQPLWPGKSDVDQLYLIRTLVETGFRHVDQAGLEL 232
Db 176 LGSRYSTPVDMMWSVGCIFAEMVNQKPLFPDSEIDELFKIFRVLGTP----- 223
Qy 233 LTSSDPPAVASQSAGITGKLIPRHQSIFKSNGFHGISIPEPEMETLEEKFSVDVHPVAL 292
Db 224 -NEQSWPGVSS-----LPDYKSAFPK-----WQAQDLATI---VPTLDPAGL 261
Qy 293 NFMKGCLXNPDRLTCSQLLESSYFDSFOEAQ 325
Db 262 DLLSKMLRYEPNKRITARQALEHEYFKDLEMVQ 294

Search completed: November 13, 2003, 14:21:59
Job time : 16.7251 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 14:16:37 ; Search time 14.6767 Seconds
(without alignments)
2273.702 Million cell updates/sec

Title: US-09-671-050-10
Perfect score: 1820
Sequence: 1 MEKYEKLAKTGGSGYGVFK.....RKARNEGRRRRQQVLPKLS 347

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1157.5	63.6	358	2 S23383	protein kinase (EC
2	1149.5	63.2	376	2 S22745	serine/threonine p
3	565	31.0	305	2 S23382	protein kinase (EC
4	563	30.9	302	1 I50474	protein kinase (EC
5	556.5	30.7	292	2 S40021	protein kinase (EC
6	555	30.5	297	2 A37671	protein kinase (EC
7	554	30.4	298	2 A41227	protein kinase (EC
8	552.5	30.4	294	2 B40444	protein kinase (EC
9	551.5	30.3	294	2 S23095	protein kinase (EC
10	551.5	30.3	294	2 S22440	protein kinase (EC
11	546.5	30.0	294	1 A40444	protein kinase (EC
12	546.5	30.0	294	2 T49271	CELL DIVISION CONT
13	545	29.9	297	2 A29539	protein kinase (EC
14	545	29.9	298	1 A44878	protein kinase (EC
15	542	29.8	297	2 I45977	cyclin-dependent k
16	541	29.7	297	1 S24913	protein kinase (EC
17	540.5	29.7	297	1 S12009	protein kinase cdc
18	540.5	29.7	302	2 B44349	protein kinase (EC
19	540	29.7	302	1 A44349	protein kinase (EC
20	539	29.6	303	1 S06012	protein kinase (EC
21	537	29.5	288	1 S42566	protein kinase (EC
22	536.5	29.5	311	2 S36619	protein kinase (EC
23	535.5	29.4	294	1 J02243	protein kinase (EC
24	535.5	29.4	294	1 S57928	protein kinase (EC
25	535.5	29.4	294	1 S42049	protein kinase (EC
26	532	29.2	301	1 S19209	protein kinase (EC
27	531	29.2	292	1 A46365	protein kinase (EC
28	530.5	29.1	294	1 S31332	protein kinase (EC
29	529	29.1	292	1 I49592	protein kinase (EC

30	529	29.1	292	1 A45091	protein kinase (EC
31	527.5	29.0	308	1 S53538	protein kinase (EC
32	527	29.0	297	1 A36074	protein kinase (EC
33	527	29.0	301	1 S42101	protein kinase (EC
34	526.5	28.9	291	2 S23386	protein kinase (EC
35	526	28.9	293	2 JE0374	cyclin-dependent k
36	526	28.9	294	2 S51008	protein kinase (EC
37	526	28.9	296	2 S24386	protein kinase (EC
38	524	28.8	302	1 OKBY85	protein kinase PHO
39	521.5	28.7	302	2 T17115	protein kinase cdc
40	519.5	28.5	346	2 AS4820	CDK-activating pro
41	518	28.5	292	2 S22441	protein kinase (EC
42	517.5	28.4	346	2 AS6231	MO15/CDK-activatin
43	516.5	28.4	314	1 S12007	protein kinase (EC
44	514	28.2	293	2 T02922	protein kinase (EC
45	514	28.2	346	1 I78840	protein kinase (EC

ALIGNMENTS

RESULT 1

S23383
protein kinase (EC 2.7.1.37) cdc2-related KXIALRE - human
C:Species: Homo sapiens (man)
C:Date: 15-Oct-1994 #sequence_revision 15-Oct-1994 #text_change 18-Jun-1999
C:Accession: S23383; S22744
R:Meyerson, M.; Enders, G.H.; Wu, C.L.; Su, L.K.; Gorka, C.; Nelson, C.; Harlow, E.; T
EMBO J. 11, 2909-2917, 1992
A:Title: A family of human cdc2-related protein kinases.
A:Reference number: S23382; MUID:92347325; PMID:1639063
A:Accession: S23383
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-358 <MEY>
A:Cross-references: EMBL:X66358; NID:g36614; PIDN:CAA47002.1; PID:g36615
C:Superfamily: Kinase-related transfor protein; protein kinase homology
C:Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
F:3-278/Domain: protein kinase homology <KIN>
F:11-19/Region: protein kinase ATP-binding motif
F:34,52,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match: 63.6%; Score 1157.5; DB 2; Length 358;
Best Local Similarity 64.7%; Pred. No. 31e-46;
Matches 211; Conservative 41; Mismatches 41; Indels 33; Gaps 2;

QY	1	MEKYEKLAKTGGSGYGVFKCRNKTSGQVAVKFFVESEDDPVVKKIALREIRMLKQLKH	60
DB	2	MEKYEKIGIGESYGVFKCRNRTDGOIVAIAKKFLESEDDPVVKKIALREIRMLKQLKH	61
QY	61	PNLVNLEVFRRKRKHVLPVEYCDHTLLNELERNPNGVADGVKSVLWQTLQALNFCIH	120
DB	62	PNLVNLEVFRRKRRLHLVFEYCDHTVLHEDRYQRGVPEHLVKSITWOTLQAVNFCXH	121
QY	121	NCIHRDIKPENILITKQGIKICDFGFAQILJ-PGDAYTDYVATRYRAPELLVGDTOYG	179
DB	122	NCIHRDVKPENILITKHSVVKLCDFGPARLLTGPSDYTYDYVATRYRSPPELLVGDTOYG	181
QY	180	SSVDIWAIGCVFAELTGGQPLWPGKSDVDQLYLIIRTLVETGFRHVDQAGLELLTSSDPP	239
DB	182	FPVDVWAIGCVFAELLSGVPLWPGKSDVDQLYLRKTL-----	219
QY	240	AVASQSAGITKLIIPHQSIFKSNGFHGISIPEPEDMETLEKPSDVHPVAFNFMKGCL	299
DB	220	-----GDLIPRQQVFTNQYFGVKIPDPEDMEPLEKLPNISPYPALGLKGL	269
QY	300	KMNPDDLRLTCSQLLESSYFDSFQEAQ	325
DB	270	HMDPTERTLTCEQLLHHPYFENIREIE	295

RESULT 2

S22745

serine/threonine protein kinase KIAJRE (EC 2.7.1.1) - human
C;Species: Homo sapiens (man)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 19-Dec-1997
C;Accession: S22745
R;Meyerson, M.L.
Submitted to the SMC Data Library, May 1992
A;Reference number: S22743
A;Accession: S22745
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-376 <MEY>
A;Cross-references: EMBL:X66359
C;Genetics:
A;Introns: 152/3; 170/3
A;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
F;3-296/Domain: protein kinase homology <KIN>
F;11-19/Region: protein kinase ATP-binding motif

Query Match 63.2%; Score 1149.5; DB 2; Length 376;
Best Local Similarity 61.3%; Pred. No. 7.5e-46;
Matches 211; Conservative 41; Mismatches 41; Indels 51; Gaps 2;

QY 1 MEKYEKLAKTGEGSYGWFKCRNKTSGOVAVKVFVESEDDPVVKKIALREIRMLKQKH 60
Db 2 MEKYEKIGKIGEGSYGWFKCRNRTGQIVAIKFESEDDPVVKKIALREIRMLKQKH 61

QY 61 PNLVNLIEVFRKRKXHLVFEYCDHTLLNELERNPENGVDGVKSVLWQTLQALNFCIH 120
Db 62 PNLVNLIEVFRKRRLHLVFEYCDHTVJHELDRYQGVPEHLVKSITWQTLQAVNFCXH 121

QY 121 NCIHRLDKPENILITKQGIKICDFGFAQIL-----PGDAYTDYV 162
Db 122 NCIHRLDKPENILITKHSVIKLCDFGFARLLFXFPQSAAVCFPCSTGTGTPSDYTDYV 182

QY 162 ATRWYRAPELLVGCTQYSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLVETG 221
Db 182 ATRWYRSELLVGDTQYGPVVDVWAIGCVFAELLSGVPLWPGKSDVDQLYLIRKTL---- 237

QY 222 FRHVDQAGJELLTSSDPPAVASQSAIGITGKLIPRHQSIFKSNFGFHGISIPEPEDMETLE 281
Db 238 -----GDLIPRHQQVFSTNQVFSGVKIPDPEDEMEPLE 269

QY 282 EKFSQVHPVALNPMKGLKMNPDRLTCSQLLESSYFDSFQEAQ 325
Db 270 LKFPNISYPALGULKGCLHMDPTERTLCQLLHHYPFENIREIE 313

RESULT 3
S23382
protein kinase (EC 2.7.1.37) cdk3 - human
C;Species: Homo sapiens (man)
C;Date: 18-Jun-1993 #sequence_revision 18-Jun-1993 #text_change 18-Jun-1999
C;Accession: S23382; S22743
R;Meyerson, M.; Enders, G.H.; Wu, C.L.; Su, L.K.; Gorka, C.; Nelson, C.; Harlow, E.; Tsai
EMBC J. 11, 2909-2917, 1992
A;Title: A family of human cdc2-related protein kinases.
A;Reference number: S23382; MUID:92347325; PMID:1639063
A;Accession: S23382
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-305 <MEY>
A;Cross-references: EMBL:X66357; NID:G36612; PIDN:CAA47001.1; PID:G36613
C;Genetics:
A;Gene: GDB:CDK3
A;Cross-references: GDB:283456
A;Map position: 12q13-12q13
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein kinase
F;2-255/Domain: protein kinase homology <KIN>
F;10-18/Region: protein kinase ATP-binding motif
F;33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 31.0%; Score 565; DB 2; Length 305;
Best Local Similarity 36.8%; Pred. No. 2.1e-19;
Matches 124; Conservative 66; Mismatches 99; Indels 48; Gaps 9;

QY 1 MEKYEKLAKTGEGSYGWFKCRNKTSGOVAVKVF---VESEDDPVVKKIALREIRMLKQ 57
Db 1 MDMFQKVEKIGEGTYGVVYKAKNRETQQLVALKIRLDLEMEGV---STAIREISLLKE 57

QY 58 LKHPNLVNLIEVFRKRKXHLVFEYCDHTLLNELERNPENG-VADGVKSVLWQTLQALNF 116
Db 58 LKHPNLVNLIEVFRKRKXHLVFEYCDHTLLNELERNPENG-VADGVKSVLWQTLQALNF 117

QY 117 CHIHNCIHRDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGD 175
Db 118 CHSHRVIHRDLKPNLLINELGAIKLADFGLARAFGVPLRTVTHEVVTLMWYRAPEILLGS 177

QY 176 TQYSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLVETGFRHVDQAGLELLTS 235
Db 178 KFYTTAVDIWSIGCIFAEVMTKALFPDSEIDQLFRIFRML----- 219

QY 236 SDPPAVASQSA-GITGKLIPRHQSIFKSNFGFHGISIPEPE-DMETLEEKFSQVHPVAL 292
Db 220 ---GTPSEDTWPGVT--QLPDYKGSF-----PKWTRKGLLEEIVPNLEPEGR 260

QY 293 NFMKGLKMNPDRLTCSQLLESSYFDSFQEAQIKRK 329
Db 261 DLEMQLLDYDPSQRITAKTALAHPYFSSPEPSPAARQ 297

RESULT 4
I50474
protein kinase (EC 2.7.1.37) cdc2 [similarity] - goldfish
C;Species: Carassius auratus (goldfish)
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 05-May-2000
C;Accession: I50474
R;Kajihara, H.; Yamashita, M.; Katsu, Y.; Nagahama, Y.
Dev. Growth Differ. 35, 647-654, 1993
A;Title: Isolation and characterization of goldfish cdc2, a catalytic component of matu
A;Reference number: I50474
A;Accession: I50474
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-302 <KAJ>
A;Cross-references: GB:D17758; NID:G471097; PIDN:BAA04605.1; PID:G471098
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: ATP; phosphotransferase
F;2-256/Domain: protein kinase homology <KIN>
F;10-18/Region: protein kinase ATP-binding motif
F;33,51,128,130/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 30.9%; Score 563; DB 1; Length 302;
Best Local Similarity 41.2%; Pred. No. 2.6e-19;
Matches 117; Conservative 60; Mismatches 71; Indels 36; Gaps 6;

QY 1 MEKYEKLAKTGEGSYGWFKCRNKTSGOVAVKVF-VESEDDPVVKKIALREIRMLKQLK 59
Db 1 MDDYLKIEKIGEGTYGVVYKGRNKTGQVAVMKKIRLESEEG-VPSTAVREISLLKELQ 59

QY 60 HPNLVNLIEVFRKRKXHLVFEYCDHTLLNELERNPENG---VADGVKSVLWQTLQALNFC 117
Db 60 HPNVVRLDLVLMQESKLYLVFEFLSMDLKKYLDSPSGQFMDPMLVKSILYQILEGILFC 119

QY 118 HIHNCIHRDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGD 176
Db 120 HCRRVJHRDLKPNLLIDNKGVIKLADFGLARAFGVPRVYVTHEVVTLMWYRAPEVLGAS 179

QY 177 QYSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTL----- 217
Db 180 RYSTPVDVMSIGTIFAELATKKPLFGDSEIDQLFRIFRITLCTPNNEVWPDVESLPDYKN 239

QY 218 -----VETGFRHVDQAGLELLTSS---DPPAVASQSAIGIT 249
Db 240 TFPKWKSGNLASTVKNLDKNGIDLLTKMLIYDPPKRISARQAWT 283

A:Molecule type: mRNA
A:Residues: 1-294 <FER>
A:Cross-references: GB:S45387; MID:g257373; PIDN:AAB23643.1; PID:g257374
A:Experimental source: flower
C:Comment: The protein is a key component of the eukaryotic cell cycle.
C:Genetics:
A:Gene: cdc2
A:Introns: 3/3; 65/2; 105/3; 163/3; 218/2; 247/3; 265/3
C:Function:
A:Description: phosphotransferase; protein kinase; required for G1 to S-phase transition
C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein kinase
F:2-256/Domain: protein kinase homology <KIN>
F:10-18/Region: protein kinase ATP-binding motif
F:33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted
Query Match 30.3%; Score 551.5; DB 2; Length 294;
Best Local Similarity 35.5%; Pred. No. 8.4e-19;
Matches 115; Conservative 71; Mismatches 95; Indels 43; Gaps 8;
Qy 1 MEKYEKLAKTGEAGYGVVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MDQYKEKIGEGTYGVVYKARDKVTNETIALKKIRLEQDEGVSPSTAIRSLLKEMQH 60
Qy 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNLELRNPNGVAD-GVIKSVLWQTLQALNFCHI 119
Db 61 SNIVKLDVHSEKRLVFEYLDLDLKKHMDSTPDSKDLHMIKTYLYQILRGIAVCHS 120
Qy 120 HNCIHRDIPENILITKQ-GIIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDQ 177
Db 121 HRVLRDLKPQNLLIDRRTNALKLADFGLARAFGIPVRTFTHEVVTLWYRAPEILLGSH 180
Qy 178 YGSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIR---TLVETGFRHVDQAGLELLT 234
Db 181 YSTPVDIWSVGCIFAEMISQKPLFPDSEIDQPKIFRIMGTPYEDTWR----- 229
Qy 235 SSDPPAVASQSAGITGKLIPRHQSIFKSNFGFHGISIPEPEMETLEKFSDVHPVALNF 294
Db 230 -----GVTS--LPDYKSAPFK-----WKPTDLET---VPNLDPDGVDL 263
Qy 295 MKGCLKMNPDRLTCSQLLESSYF 318
Db 264 LSKMLMDPTKRINARAALAEHYF 287
RESULT 10
S22440
protein kinase (EC 2.7.1.37) cdc2 homolog 1 - rice
C:Species: Oryza sativa (rice)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 18-Jun-1999
C:Accession: S22440
R:Hashimoto, J.; Hirabayashi, T.; Hayano, Y.; Hata, S.; Chashi, Y.; Suzuki, I.; Utsugi, M.; Gen. Genet. 233, 10-16, 1992
A:Title: Isolation and characterization of cDNA clones encoding cdc2 homologues from Ory
A:Reference number: S22440; MUID:92293101; PMID:1376401
A:Accession: S22440
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-294 <HAS>
A:Cross-references: EMBL:X60374; MID:g20342; PIDN:CAA42922.1; PID:g20343
C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein kinase
F:2-256/Domain: protein kinase homology <KIN>
F:10-18/Region: protein kinase ATP-binding motif
F:33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted
Query Match 30.3%; Score 551.5; DB 2; Length 294;
Best Local Similarity 34.8%; Pred. No. 8.4e-19;
Matches 116; Conservative 68; Mismatches 102; Indels 47; Gaps 8;
Qy 1 MEKYEKLAKTGEAGYGVVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MEQYEKEKIGEGTYGVVYKARDKVTNETIALKKIRLEQDEGVSPSTAIRSLLKEMQH 60

Qy 61 PNLVNLIEVFRKRKMHVFEYCDHTLLN-----ELERNPNGVADGVIKSVLWQTLQAL 114
Db 61 GNVRLHDVHSEKRIYLVFEYLDLDLKKEMDSCPEFAKNPT-----LIKSYLYQILRGV 115
Qy 115 NFCHIHNCIHRDIPENILITKQ-GIIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELL 172
Db 116 AYCHSHRVLHRDLKPQNLLIDRRTNALKLADFGLARAFGIPVRTFTHEVVTLWYRAPEIL 175
Qy 173 VGDTQYSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLVETGFRHVDQAGLEL 232
Db 176 LGSRYSTPVDVMSVGCIFAEMVNQKPLFPDSEIDELFKIFRVLGTP----- 223
Qy 233 LTSSDPPAVASQSAGITGKLIPRHQSIFKSNFGFHGISIPEPEMETLEKFSDVHPVAL 292
Db 224 -NEQSWPGVSS-----LPDYKSAPFK-----WQAQDLATI---VPTLDPAGL 261
Qy 293 NFMKGCLKMNPDRLTCSQLLESSYFDSFQEAQ 325
Db 262 DLLSKMLRYEPNKRITARQALEHEYFKDLEVMVQ 294
RESULT 11
A40444
protein kinase (EC 2.7.1.37) cdc2 homolog A - maize
C:Species: Zea mays (maize)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A40444
R:Colasanti, J.; Tyers, M.; Sundaresan, V.
Proc. Natl. Acad. Sci. U.S.A. 88, 3377-3381, 1991
A:Title: Isolation and characterization of cDNA clones encoding a functional p34 (cdc2)
A:Reference number: A40444; MUID:91195354; PMID:2014258
A:Accession: A40444
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-294 <COL>
A:Cross-references: GB:M60526
C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; phosphotransferase
F:2-256/Domain: protein kinase homology <KIN>
F:10-18/Region: protein kinase ATP-binding motif
F:33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted
Query Match 30.0%; Score 546.5; DB 1; Length 294;
Best Local Similarity 34.1%; Pred. No. 1.4e-18;
Matches 117; Conservative 66; Mismatches 93; Indels 67; Gaps 9;
Qy 1 MEKYEKLAKTGEAGYGVVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MEQYEKEKIGEGTYGVVYKALDKTANETIALKKIRLEQDEGVSPSTAIRSLLKEMNH 60
Qy 61 PNLVNLIEVFRKRKMHVFEYCDHTLLN-----ELERNPNGVADGVIKSVLWQTLQAL 114
Db 61 GNVRLHDVHSEKRIYLVFEYLDLDLKKPMDSCEFAKNPT-----LIKSYLYQILHGV 115
Qy 115 NFCHIHNCIHRDIPENILITKQ-GIIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELL 172
Db 116 AYCHSHRVLHRDLKPQNLLIDRRTNALKLADFGLARAFGIPVRTFTHEVVTLWYRAPEIL 175
Qy 173 VGDTQYSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLVETGFRHVDQAGLEL 232
Db 176 LGARQYSTPVDVMSVGCIFAEMVNQKPLFPDSEIDELFKIFRILG----- 221
Qy 233 LTSSDPPAVASQSAGITGKLIPRHQSIFKSNFGFHGIS-IP-----EPEDMETLEE 282
Db 222 -----TPNEQS-----WPGVSCLPDFKTAFFRWQAQDLATV-- 252
Qy 283 KFSDVHPVALNFMKGCLKMNPDRLTCSQLLESSYFDSFQEAQ 325
Db 253 -VPNLDPAGLDLLSKMLRYEPEPKRITARQALEHEYFKOLEVVQ 294
RESULT 12

Search completed: November 13, 2003, 14:21:16
Job time : 15.6767 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 14:12:56 ; Search time 10.4934 Seconds
(without alignments)
1556.584 Million cell updates/sec

Title: US-09-671-050-10
Perfect score: 1820
Sequence: 1 MEKYEKLAKTGGSGVGVFK.....RKARNEGRRRRQQVLEPKS 347

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1157.5	63.6	358	1	Q00532 homo sapien
2	694.5	38.2	1030	1	Q76039 homo sapien
3	565	31.0	305	1	Q00526 homo sapien
4	563	30.9	302	1	P51958 carassius a
5	561	30.8	297	1	P23437 xenopus lae
6	558.5	30.7	292	1	P34117 dictyosteli
7	554	30.4	298	1	P24941 homo sapien
8	551.5	30.3	294	1	P29618 oryza sativ
9	551.5	30.3	294	1	P24102 arabidopsis
10	550	30.2	298	1	O55076 cricetus
11	549.5	30.2	303	1	Q9dga2 oryzias jav
12	548	30.1	298	1	Q63699 rattus norv
13	546	30.0	298	1	P48963 mesocricetu
14	545	29.9	297	1	P06493 homo sapien
15	545	29.9	298	1	P43450 carassius a
16	543.5	29.9	303	1	Q9dgd3 oryzias lat
17	542.5	29.8	294	1	P93101 chenopodium
18	542	29.8	297	1	P48734 bos taurus
19	541.5	29.8	294	1	P23111 zea mays (m
20	541.5	29.8	303	1	Q9dga5 oryzias cur
21	541.5	29.8	303	1	Q9dgg98 oryzias luz
22	541	29.7	297	1	P39951 rattus norv
23	540.5	29.7	297	1	P23572 drosophila
24	540.5	29.7	302	1	P24033 xenopus lae
25	540	29.7	302	1	P35567 xenopus lae
26	539	29.6	303	1	P13863 gallus gall
27	537	29.5	288	1	Q07785 plasmodium
28	536.5	29.5	294	1	P52389 vigna ungu
29	536.5	29.5	311	1	P54656 trypanosoma
30	535.5	29.4	294	1	Q41639 vigna accni
31	532	29.2	297	1	P1440 mus musculu
32	532	29.2	301	1	P18973 trypanosoma
33	531	29.2	292	1	Q02399 bos taurus

RESULT 1
KXIA_HUMAN
ID KXIA_HUMAN STANDARD; PRT; 358 AA.
AC Q00532;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Serine/threonine-protein kinase KXIALRE (EC 2.7.1.-) (Cyclin-dependent
DE kinase-like 1).
GN CDKL1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92347325; PubMed=1639063;
RA Meyerson M., Enders G.H., Wu C.-L., Su L.-K., Gorka C., Nelson C.,
RA Harlow E., Tsai L.-H.;
RT "A family of human cdc2-related protein kinases.";
RL EMBO J. 11:2909-2917(1992).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; X66358; CAA47002.1; --
EMBL; X66359; CAA47002.1; JOINED.
PIR; S23383; S23383.
HSSP; P24941; LHCL.
Genew; HGNC:1781; CDKL1.
GK; Q00532; --
VIM; 603441; --
GO; GO:0004693; F:cyclin-dependent protein kinase activity; TAS.
GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
GO; GO:0000074; P:regulation of cell cycle; TAS.
InterPro; IPR000719; Prot_kinase.
InterPro; IPR002290; Ser_thr_kinase.
Pfam; PF00069; pkinase; 1.
ProDom; PD000001; Prot_kinase; 1.
SMART; SM00220; S_TKC; 1.
PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
Transferase; Serine/threonine-protein kinase; ATP-binding.
DOMAIN 5 288 PROTEIN_KINASE.
NP_BIND 11 19 ATP (BY SIMILARITY).
BINDING 34 34 ATP (BY SIMILARITY).
ACT_SITE 127 127 BY SIMILARITY.
SEQUENCE 358 AA; 41834 MW; 88344321F24B77C6 CRC64;

ALIGNMENTS

34	531	29.2	292	1	CDK5 RAT	Q03114 rattus norv
35	530.5	29.1	294	1	CC22_MEDSA	Q05C06 medicago sa
36	529	29.1	292	1	CDK5_MOUSE	P49615 mus musculu
37	527	29.0	301	1	CC21_TRYCO	P54664 trypanosoma
38	526	28.9	292	1	CDK5_HUMAN	Q00535 homo sapien
39	526	28.9	296	1	CDK2_DICDI	P34112 dictyosteli
40	525	28.8	294	1	CDK5_DROME	P48609 drosophila
41	521.5	28.7	294	1	CC2A_ANTMA	Q38772 antirrhinum
42	521	28.6	292	1	CDK5_XENLA	P51166 xenopus lae
43	519.5	28.5	346	1	CDK7_HUMAN	P50613 homo sapien
44	518	28.5	292	1	CC22_ORYSA	P29619 oryza sativ
45	518	28.5	305	1	PH85_YEAST	P17157 saccharomyc

```
Query Match      63.6%; Score 1157.5; DB 1; Length 358;
Best Local Similarity 64.7%; Pred. No. 9.8e-75;
Matches 211; Conservative 41; Mismatches 41; Indels 33; Gaps 2;

QY 1 MEKYEKLAKTGEAGSYGVVFKCRNKTSGQVAVKVFSEDDPVVKIALREIRMLKQLKH 60
DB 2 MEKYEKIGKIGESYGVVFKCRNRTDQGVAVKVFSEDDPVVKIALREIRMLKQLKH 61
QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEPNPVGADGVKSVLWQTLQALNFCIH 120
DB 62 PNLVNLIEVFRKRRLHLVFEYCDHTVLHELDYQGVPEHLVKSITWQTLQAVNFCIH 121
QY 121 NCIHRIKPNILITKQGIKICDFGFAQILI-PGDAYTDYVATRWYRAPELLVGDQY 179
DB 122 NCIHRIKPNILITKHSVIKLCDFGARLLTGPSDYTDYVATRWYRSPPELLVGDQY 181
QY 180 SSVDIWAIGCVFAELLTGQPLWPGKSDVDQVLIIRTLVETGFRHVDQAGLELLTSSDP 239
DB 182 PVDVWAIGCVFAELLTGQPLWPGKSDVDQVLIIRTLVETGFRHVDQAGLELLTSSDP 239
QY 240 AVASQSAGITGKLIPRHQSIKSNFGEHGISPEPEMETLEEKFSQVHPVAFNFKGCL 299
DB 220 -----GDLIPRHQCVFSTNCFYSGVKIPDPEDMEFELKAFNISYPALGCLKGCL 269
QY 300 KMYPDDRLTCSQLLESSYFDSFQEAQ 325
DB 270 HMDPTERTCEQLLHPYFENIREIE 295

RESULT 2
STK9_HUMAN STANDARD; PRT; 1030 AA.
ID STK9_HUMAN
AC O76039; Q14198;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Serine/threonine protein kinase 9 (EC 2.7.1.37).
GN STK9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98389628; PubMed=9721213;
RA Montini E., Andolfi G., Caruso A., Buchner G., Walpole S.M.,
RA Mariani M., Consalez G.G., Trump D., Ballabio A., Franco B.;
RT "Identification and characterization of a novel serine-threonine
RT kinase gene from the Xp22 region.";
RL Genomics 51:427-433(1998).
RN [2]
RP SEQUENCE OF 339-789 FROM N.A.
RA Krause S.W., Rehli M., Kreutz M., Schwarzfischer L., Paulauskis J.E.,
RA Andreesen J.D.;
RT "Differential screening leads to novel genetic markers of monocyte
RT to macrophage maturation.";
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- ENZYME REGULATION: COULD BE ACTIVATED BY PHOSPHORYLATION ON
CC TYROSINE AND THREONINE. PROBABLY BOTH PHOSPHORYLATIONS ARE
CC REQUIRED FOR ACTIVITY.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN BRAIN, LUNG, KIDNEY, PROSTATE,
CC OVARY, PLACENTA, PANCREAS AND TESTIS.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.
CC -!- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-10 IS THE INITIATOR.
CC -!- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC FRAMESHIFT IN POSITION 415.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```

```
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y15057; CAA75342.1; -.
DR EMBL; X89059; CAA61445.1; ALT_FRAME.
DR HSSP; P24941; 1B38.
DR Genew; HGNC:11411; STK9.
DR GK; O76039; -.
DR MIM; 300203; -.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; TAS.
DR GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation.
FT DOMAIN 13 297 PROTEIN KINASE.
FT NP_BIND 19 27 ATP (BY SIMILARITY).
FT BINDING 42 42 ATP (BY SIMILARITY).
FT ACT_SITE 135 135 BY SIMILARITY.
FT MOD_RES 169 169 PHOSPHORYLATION (ACTIVATES THE KINASE)
FT (POTENTIAL).
FT MOD_RES 171 171 PHOSPHORYLATION (ACTIVATES THE KINASE)
FT (POTENTIAL).
FT DOMAIN 784 789 POLY-LYS.
FT CONFLICT 339 340 HR -> GT (IN REF. 2).
FT CONFLICT 541 541 L -> W (IN REF. 2).
FT CONFLICT 731 764 MISSING (IN REF. 2).
SQ SEQUENCE 1030 AA; 115537 MW; 8A1C9C438610EF08 CRC64;

Query Match      38.2%; Score 694.5; DB 1; Length 1030;
Best Local Similarity 40.3%; Pred. No. 1.5e-41;
Matches 139; Conservative 69; Mismatches 98; Indels 39; Gaps 5;

QY 1 MEKYEKLAKTGEAGSYGVVFKCRNKTSGQVAVKVFSEDDPVVKIALREIRMLKQLKH 60
DB 10 MNKFEILGVGEGAYGVVLCRHKETHEIVAIFKFDSENEEVKETTRELKYLRLTKQ 69
QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEPNPVGADGVKSVLWQTLQALNFCIH 120
DB 70 ENIVELKEAFRRRCKLYLVFEYVEKKNMELLEEMPNGVPPKVKSYIQLIKAIHWCHKN 129
QY 121 NCIHRIKPNILITKQGIKICDFGFAQILI-PGD--AYTDYVATRWYRAPELLVGDQY 178
DB 130 DIVHRDIKPNILISHNDVLKLCDFGARNLSEGNANNTYEVATRWYRSPPELLLG-APY 188
QY 179 GSSVDIWAIGCVFAELLTGQPLWPGKSDVDQVLIIRTLVETGFRHVDQAGLELLTSSDP 238
DB 189 GKSVDMMWSVGCILGELSDGQPLFFGESEIDQLFTIQVL----- 227
QY 239 PAVASQSAGITGKLIPRHQSIKSNFGEHGISPEPEMETLEEKFSQVHPVAFNFKMG 297
DB 228 -----GPLPSEQMKLFYSNPRFRHGLRFPVAVNHPQSLERRYLGIILSVLLDMKN 276
QY 298 CLKMNPDDRLTCSQLLESSYFDSFQEAQIKRKARNEGRRRRQQV 342
DB 277 LLKLDPADRYLTEQCLN---HPTFTQRLDRLDRSPRSRAKRPYHV 318

RESULT 3
CDK3_HUMAN STANDARD; PRT; 305 AA.
ID CDK3_HUMAN
AC Q00526;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
```


RT "The isolation from a unicellular organism, Dictyostelium discoideum,
of a highly-related cdc2 gene with characteristics of the PCTAIRE
subfamily."
RL Biochim. Biophys. Acta 1179:117-124(1993).
CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
CC THE ENZYME (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L00652; AAA16056.1; -.
DR PIR; S40021; S40021.
DR HSSP; P24941; 1CKP.
DR DictyDb; D05039; crpA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC_1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
KW transferase; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation.
FT DOMAIN 4 285 PROTEIN KINASE.
FT NP_BIND 10 18 ATP (BY SIMILARITY).
FT BINDING 33 33 ATP (BY SIMILARITY).
FT ACT_SITE 126 126 BY SIMILARITY.
FT MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 292 AA; 33226 MW; 361AB54C4E3BD41E CRC64;

Query Match 30.7%; Score 558.5; DB 1; Length 292;
Best Local Similarity 35.0%; Pred. No. 1.4e-32;
Matches 115; Conservative 66; Mismatches 10; Indels 47; Gaps 6;

QY 1 MEKYEKLAKTGEISYGVVFXCRNKTSGQWAVKRFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MEKYSKEKLGEGYGIIVNKAQKQKRETEIIVKLRIRLDESEDEGVPCPTAIRE-SLLKELKH 60

QY 61 PNLVNLIEVFERKRMHLVFEYCDHTLLNELEARNPNGVADGVIKSVLWQTLQALNFCIH 120
Db 61 PNIIVRHQVHTERKLTIVFEYLDQDLKKYLDCEGGGEISKPTIKSFYQLKGVAFCHDH 120

QY 121 NCIHRIKPENILITKQIIKICDFGFAQIL-IPGDAYTDYVATRWYRAPPELLVGETQYG 179
Db 121 RVLHRDLKPQNLINRKGLKLADEGLARAFGIPVITYSHEWVTLWYRAPDVLMGSRKYS 180

QY 180 SSVDIWAIGCVFAELLTGQPLWPKSDVDQLYLIRTLVETGFRHVQDQAGLELLTSSDPP 239
Db 181 TPIDIWSALCIFAEMASGRPLFGSGTSDQLFRIFKLGTP-----NEESWP 227

QY 240 AVASQSAGITGKLIPRHQSIF-----KSNGFHGISPEPEDMETLEEFSDVHPVALN 293
Db 228 SITE-----LPEYKTDFPVHPAHQLSSIVHG-----LDEK-----GLN 260

QY 294 FMKGLKKNPDDRLTCSQLLESSYFDSFQ 322
Db 261 LLSKMLQVDPNQIRITAAALKHPYFDGLE 289

RESULT 7
CDK2 HUMAN STANDARD; PRT: 298 AA.
AC P24941;
DT 01-MAR-1992 (Rel. 21, Created)

DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Cell division protein kinase 2 (EC 2.7.1.1-) (p33 protein kinase).
GN CDK2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91330891; PubMed=1714386;
RA Elledge S.J., Spottswood M.R.;
RT "A new human p34 protein kinase, CDK2, identified by complementation
RT of a cdc28 mutation in Saccharomyces cerevisiae, is a homolog of
RT Xenopus Egl.";
RL EMBO J. 10:2653-2659(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91367262; PubMed=1653904;
RA Tsai L.-H., Harlow E., Meyerson M.;
RT "Isolation of the human cdk2 gene that encodes the cyclin A- and
RT adenovirus E1A-associated p33 kinase.";
RL Nature 353:174-177(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92020980; PubMed=1717994;
RA Ninomiya-Tsuji J., Nomoto S., Yasuda H., Reed S.I., Matsumoto K.;
RT "Cloning of a human cDNA encoding a CDC2-related kinase by
RT complementation of a budding yeast cdc28 mutation.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:9006-9010(1991).
RN [4]
RP SEQUENCE FROM N.A.
RA Rieder M.J., Livingston R.J., Braun A.C., Montoya M.A., Chung M.-W.,
RA Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,
RA Schackwitz W.S., Sherwood J.K., Witrak L.A., Nickerson D.A.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Shevchenko Y., Bouffard A.,
RA Whiting M., Madan A., Young A.C., Green E.D., Dickson M.C.,
RA Blakesley R.W., Touchman J.W., Schmutz J., Myers R.M.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Smailus D.E.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Marra M.A.;
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [6]
RP PHOSPHORYLATION SITES.
RX MEDLINE=93010995; PubMed=1396589;
RA Gu Y., Rosenblatt J., O'Morgan D.O.;
RT "Cell cycle regulation of CDK2 activity by phosphorylation of Thr160
RT and Tyr15.";
RL EMBO J. 11:3995-4005(1992).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
RX MEDLINE=93288132; PubMed=8510751;
RA de Bondt H.L., Rosenblatt J., Jancarik J., Jones H.D.,
RA Morgan D.O., Kim S.-H.;
RT "Crystal structure of cyclin-dependent kinase 2.";

```
RL Nature 363:595-602(1993).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF COMPLEX WITH CYCLIN A.
RX MEDLINE=95356811; PubMed=7630397;
RA Jeffrey P.D., Russo A.A., Polyak K., Gibbs E., Hurwitz J.,
RA Massague J., Pavletich N.P.;
RT "Mechanism of CDK activation revealed by the structure of a
RL cyclinA-CDK2 complex.";
RN [9]
RP X-RAY CRYSTALLOGRAPHY (2.33 ANGSTROMS) OF COMPLEX WITH L868276.
RX MEDLINE=96181476; PubMed=8610113;
RA de Azevedo W.F. Jr., Muleer-Dieckmann H.-J., Schulze-Gahmen U.,
RA Worland P.J., Sausville E., Kim S.-H.;
RT "Structural basis for specificity and potency of a flavonoid
RL inhibitor of human CDK2, a cell cycle kinase.";
RN [10]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF COMPLEX WITH CG2A AND K1P1.
RX MEDLINE=96300318; PubMed=8684460;
RA Russo A.A., Jeffrey P.D., Patten A.K., Massague J., Pavletich N.P.;
RT "Crystal structure of the p27Kip1 cyclin-dependent-kinase inhibitor
RL bound to the cyclin A-Cdk2 complex.";
RN [11]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF COMPLEX WITH CG2A.
RX MEDLINE=96313126; PubMed=8756328;
RA Russo A.A., Jeffrey P.D., Pavletich N.P.;
RT "Structural basis of cyclin-dependent kinase activation by
RL phosphorylation.";
RN [12]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE=97075215; PubMed=8917641;
RA Schulze-Gahmen U., de Bondt H.L., Kim S.-H.;
RT "High-resolution crystal structures of human cyclin-dependent kinase
RL 2 with and without ATP: bound waters and natural ligand as guides for
RL inhibitor design.";
RN [13]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=97475219; PubMed=9334743;
RA Lawrie A.M., Noble M.E.M., Tunnah P., Brown N.R., Johnson L.N.,
RA Erdicott J.A.;
RT "Protein kinase inhibition by staurosporine revealed in details of
RL the molecular interaction with CDK2.";
RN [14]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF COMPLEX WITH CKS1.
RX MEDLINE=96182647; PubMed=8601310;
RA Bourne Y., Watson M.H., Hickey M.J., Holmes W., Rocque W., Reed S.I.,
RA Tainer J.A.;
RT "Crystal structure and mutational analysis of the human CDK2 kinase
RL complex with cell cycle-regulatory protein CksHs1.";
RN [15]
RP X-RAY CRYSTALLOGRAPHY (2.05 ANGSTROMS).
RX MEDLINE=98342369; PubMed=9677190;
RA Gray N.S., Wodicka L., Thunnissen A.-M.W.H., Norman T.C., Kwon S.,
RA Espinoza F.H., Morgan D.O., Barnes G., Leclerc S., Meijer L.,
RA Kim S.H., Lockhart D.J., Schultz P.G.;
RT "Expliciting chemical libraries, structure, and genomics in the search
RL for kinase inhibitors.";
RN [16]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=981533538(1998).
RA [16]
CC -!- FUNCTION: PROBABLY INVOLVED IN THE CONTROL OF THE CELL CYCLE.
CC INTERACTS WITH CYCLINS A, D, OR E. ACTIVITY OF CDK2 IS MAXIMAL
CC DURING S PHASE AND G2.
CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
CC THE ENZYME, WHILE PHOSPHORYLATION AT THR-160 ACTIVATES IT.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X61622; CAA43807.1; -
DR EMBL; X62071; CAA43985.1; -
DR EMBL; M68520; AAA35667.1; -
DR EMBL; AF512553; AAM34794.1; -
DR EMBL; BC003065; AAH03065.1; -
DR PIR; A41227; A41227.
DR PDB; 1FIN; 27-JAN-97.
DR PDB; 1HCK; 07-DEC-96.
DR PDB; 1HCL; 07-DEC-96.
DR PDB; 1AQ1; 12-NOV-97.
DR PDB; 1JST; 11-JAN-97.
DR PDB; 1JSU; 29-JUL-97.
DR PDB; 1BUH; 09-SEP-98.
DR PDB; 1B38; 23-DEC-98.
DR PDB; 1B39; 23-DEC-98.
DR PDB; 1CKP; 13-JAN-99.
DR PDB; 1D18; 28-JAN-03.
DR PDB; 1DM2; 31-MAY-00.
DR PDB; 1E1V; 10-MAY-01.
DR PDB; 1E1X; 10-MAY-01.
DR PDB; 1F5Q; 27-DEC-00.
DR PDB; 1FQ1; 09-MAY-01.
DR PDB; 1FVT; 17-JAN-01.
DR PDB; 1FWV; 17-JAN-01.
DR PDB; 1G5S; 14-NOV-01.
DR PDB; 1GIH; 06-FEB-02.
DR PDB; 1GI1; 06-NOV-02.
DR PDB; 1GIJ; 06-FEB-02.
DR PDB; 1GY3; 06-JUN-02.
DR PDB; 1H1P; 19-SEP-02.
DR PDB; 1H1Q; 19-SEP-02.
DR PDB; 1H1R; 19-SEP-02.
DR PDB; 1H1S; 19-SEP-02.
DR PDB; 1H24; 01-FEB-03.
DR PDB; 1H25; 01-FEB-03.
DR PDB; 1H26; 01-FEB-03.
DR PDB; 1H28; 01-FEB-03.
DR PDB; 1JSV; 29-AUG-01.
DR PDB; 1JVP; 21-DEC-01.
DR PDB; 1KE5; 14-MAY-02.
DR PDB; 1KE6; 14-MAY-02.
DR PDB; 1KE7; 14-MAY-02.
DR PDB; 1KE8; 14-MAY-02.
DR PDB; 1KE9; 14-MAY-02.
Query Match 30.4%; Score 554; DB 1; Length 298;
Best Local Similarity 38.4%; Pred No. 2.9e-32;
Matches 118; Conservative 54; Mismatches 89; Indels 46; Gaps 6;
Qy 1 MEKYEKLAKTGECSYGVVFKCRNKTSGQVAVKKFVSEDDPVVKKIALREIRMLKQLKH 60
Db 1 MENFQKVEKIGEGTYGVVYKARNKLTGEWALKIRLDTETEGVPSTAIRISLLKELNH 60
Qy 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELERNP-NGVADGVIKSVLWQTLQALNFCIH 119
Db 61 PNIIVKLLDVIHTENKLYLVFEFLHQDLKFKFMDASALTGIPLPKLSYLFQLQLGAFCHS 120
Qy 120 HNCIHRDIKPENILITKQIHKICDPGFAQIL-IPGDAYTDYVATRWYRAPELLVGDQY 178
Db 121 HRVLHRLDKPQNLINTEGAIKLADFGLARAFGVPRVTYTHEVTLWYRAPEILLGCKYV 180
Qy 179 GSSVDIWAIGCVFAELLTGQPLWPKSDVDQLYLIIRTL----- 217
Db 181 STAVDIWSLGCIFAEMVTRRALFPGDSEIDQLFRITLGTDPDEVWPGVTSMPDYKPSF 240
Qy 218 ---VETGFRHV---DQAGLELLTSS---DPPAVASQSACITGKLIPRHQSIFKSNQFFH 267
```



```
Db 241 PKWQRQDFSKVPLDDEGRSLLSQMLHYDPNKRISAKAAL-----AHPPFQ 287
Qy 268 GISIEEP 274
Db 288 DVTKPVP 294

RESULT 8
CC2: ORYSA
ID CC21 ORYSA STANDARD; PRT; 294 AA.
AC P29618;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Cell division control protein 2 homolog 1 (EC 2.7.1.-).
GN CDC2-1.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Cryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RX MEDLINE=92293101; PubMed=1376401;
RA Hashimoto J., Hirabayashi T., Hayano Y., Hata S., Ohashi Y.,
RA Suzuki I., Utsugi T., Toh-E A., Kikuchi Y.;
RT "Isolation and characterization of cDNA clones encoding cdc2
RT homologues from Oryza sativa: a functional homologue and cognate
RT variants.";
RL Mol. Gen. Genet. 233:10-16(1992).
CC -!- FUNCTION: PLAYS A KEY ROLE IN THE CONTROL OF THE EUKARYOTIC
CC CELL CYCLE. COMPONENT OF THE KINASE COMPLEX THAT PHOSPHORYLATES
CC THE REPETITIVE CARBOXYL-TERMINUS OF RNA POLYMERASE II.
CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
CC THE ENZYME, WHILE PHOSPHORYLATION AT THR-161 ACTIVATES IT (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X60374; CAA42922.1; -
CC F1R; S22440; S22440.
CC HSSP; P24941; IHCL.
CC Gramene; P29618; -.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR002290; Ser_thr_kinase.
CC Pfam; PF00069; pkinase; 1.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00220; S_TKc; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
CC Transferase; Serine/threonine-protein kinase; ATP-binding;
CC Cell cycle; Cell division; Mitosis; Phosphorylation.
CC DOMAIN 4 287 PROTEIN KINASE.
FT NP BIND 10 18 ATP (BY SIMILARITY).
FT BINDING 33 33 ATP (BY SIMILARITY).
FT ACT_SITE 127 127 BY SIMILARITY.
FT MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 161 161 PHOSPHORYLATION (BY CAK) (BY SIMILARITY).
SQ SEQUENCE 294 AA; 34071 MW; 51322D93AEF4C131 CRC64;
```

Query Match 30.3%; Score 551.5; DB 1; Length 294;

```
Best Local Similarity 34.8%; Pred. No. 4.3e-32;
Matches 116; Conservative 68; Mismatches 102; Indels 47; Gaps 8;

Qy 1 MEKYEKLAKTGEGSYGVVFKCRNKTSGQVAVVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MEQYEKEKIGEGTYGVVYRARDKVTNETIALKKIRLEQDEGVFSTAIRISLLKEMHH 60

Qy 61 PNLVNLIEVFRKRKQHLVFEYCDHTLLN-----ELERNPVGADGVKSVLMQTLQAL 114
Db 61 GNIVRLHDVTHSEKRIYLVFEYLDLKKFMDSCPEFAKNPT-----LIKSYLQILRGV 115

Qy 115 NPSCHIHNCIHRDKPENILITKQ-GIKICDFGFAQIL-IPGDYTDYVATRWYRAPELL 172
Db 115 AYCHSHRVLHRLDKPQNLIDRRTNALKLADFGLARAFGIPVTRTFHEVVTWYRAPEIL 175

Qy 173 VGDTOYGSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLVETGFRHVDCAGLEL 232
Db 176 LGSRQYSTPVDMMWSVGCIFAEMVNQKPLFPDSEIDELFKIFRVLGTP----- 223

Qy 233 LTSSDPPAVASQSGITGKLIPRHQSIFKSNNGFFHGISIPEPEDMETLEKFSVDVHPVAL 292
Db 224 -NEQSWPGVSS-----LPQYKSAFPK-----WQAQDLATI---VPTLDPAGL 261

Qy 293 NPYKGLKMPNDDRLTCSQLLESSYFDSFOEAQ 325
Db 262 DLLSKMLRYEPNKRITARQALEHEYFKDLEMVQ 294

RESULT 9
CC2A_ARATH STANDARD; PRT; 294 AA.
ID CC2A_ARATH STANDARD; PRT; 294 AA.
AC P24100; Q9M307;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cell division control protein 2 homolog A (EC 2.7.1.-).
GN CDC2A OR CDC2 OR AT3G48750 OR T21J18.20.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93005715; PubMed=1840925;
RA Ferreira P.C.G., Hemerly A.S., Villarroel R., van Montagu M., Inze D.;
RT "The Arabidopsis functional homolog of the p34cdc2 protein kinase.";
RL Plant Cell 3:531-540(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=92039027; PubMed=1937013;
RA Hirayama T., Imajuku Y., Anai T., Matsui M., Oka A.;
RT "Identification of two cell-cycle-controlling cdc2 gene homologs in
RT Arabidopsis thaliana.";
RL Gene 105:159-165(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92316202; PubMed=1618302;
RA Imajuku Y., Hirayama T., Endoh H., Oka A.;
RT "Exon-intron organization of the Arabidopsis thaliana protein kinase
RT genes CDC2a and CDC2b.";
RL FEBS Lett. 304:73-77(1992).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=92339744; PubMed=1634002;
RA Inze D., Ferreira P.C.G., Hemerly A.S., van Montagu M.;
RT "Control of cell division in plants.";
RL Biochem. Soc. Trans. 20:80-84(1992).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016720; PubMed=11130713;
```



```
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ223949; CAA11680.1; -.
DR HSSP; P24941; 1A01.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding; Cell cycle;
KW Cell division; Mitosis; Phosphorylation; Alternative splicing.
FT DOMAIN 4 286
FT NP_BIND 10 18 ATP (BY SIMILARITY).
FT BINDING 33 33 ATP (BY SIMILARITY).
FT ACT_SITE 127 127 BY SIMILARITY.
FT MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 160 160 PHOSPHORYLATION (BY CAK) (BY SIMILARITY).
SQ SEQUENCE 298 AA; 33873 MW; 3B580D8C2460CD98 CRC64;

Query Match 30.2%; Score 550; DB 1; Length 298;
Best Local Similarity 38.1%; Pred. No. 5.6e-32;
Matches 117; Conservative 55; Mismatches 89; Indels 46; Gaps 6;

QY 1 MEKYEKLAKTGEISYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
DB 1 MENFQKVEKIGEGTYGVVYKAKNKLGTGEVVALKXIRLDTETEGVPSTAIRISLLKELNH 60
QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELERNP-NGVADGVKSVLWQTLQALNFC 119
DB 61 PNIIVKLLDVIHTENKLYLVFEFLHQDLKKFMDASAVTGIPPLIKSYLFQLLQGLAFCHS 120
QY 120 HNCIHRDIKPNENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDTQY 178
DB 121 HRVLHRDLKPQNLLINAEKSIKLADEGLARAFGVPRVTYTHEVVTJWYRAPEILLGCKYY 180
QY 179 GSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIRTL----- 217
DB 181 STAVDIWSLGCIFAEMVTRRALPPGDSEIDQLFRIFRTLGTDPDEVVWPGVTMPDYKPSF 240
QY 218 ---VETGFRHV---DQAGLELLTSS---DPPAVASQSAGITGKLIPRHQSIFKSNQFFH 267
DB 241 PKWARQDFSKVVPPLDGDGRSLLSQMLHYDPNKRISAKAAL-----AHQFFQ 267
QY 268 GISIPEP 274
DB 288 DVTKPVP 294
```

```
RESULT 11
CDC2_ORYJA STANDARD; PRT; 303 AA.
AC Q9DGA2; Q9DGA1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cell division control protein 2 homolog (EC 2.7.1.-) (p34 protein
DE kinase) (Cyclin-dependent kinase 1) (CDK1).
GN CDC2.
OS Oryzias javanicus (Japanese ricefish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
OX NCBI_TaxID=123683;
RN [?]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RA Yamashita M., Mita K.;
RT "cDNA cloning of Cdc2 and cyclin B in medaka species.";
```

```
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: PLAYS A KEY ROLE IN THE CONTROL OF THE EUKARYOTIC CELL
CC CYCLE. IT IS REQUIRED IN HIGHER CELLS FOR ENTRY INTO S-PHASE AND
CC MITOSIS. COMPONENT OF THE KINASE COMPLEX THAT PHOSPHORYLATES THE
CC REPETITIVE CARBOXYL-TERMINUS OF RNA POLYMERASE II. CATALYTIC
CC COMPONENT OF MPF (By similarity).
CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
CC THE ENZYME, WHILE PHOSPHORYLATION AT THR-161 ACTIVATES IT (BY
CC SIMILARITY).
CC -!- SUBUNIT: FORMS A STABLE BUT NON-COVALENT COMPLEX WITH CYCLIN B IN
CC MATURE COCYTES (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDX SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB050461; BAB17219.1; -.
DR EMBL; AB050462; BAB17220.1; -.
DR HSSP; Q00534; 1BI8.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Cell cycle; Cell division; Mitosis; Nuclear protein; Phosphorylation.
FT DOMAIN 4 287
FT NP_BIND 10 18 ATP (BY SIMILARITY).
FT BINDING 33 33 ATP (BY SIMILARITY).
FT ACT_SITE 128 128 BY SIMILARITY.
FT MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 161 161 PHOSPHORYLATION (BY CAK) (BY SIMILARITY).
FT VARIANT 49 49 V -> I.
FT VARIANT 245 245 M -> K.
FT VARIANT 269 269 I -> T.
FT VARIANT 276 276 I -> V.
FT VARIANT 299 299 C -> S.
SQ SEQUENCE 303 AA; 34691 MW; 439B8B012BE28D9C CRC64;

Query Match 30.2%; Score 549.5; DB 1; Length 303;
Best Local Similarity 35.7%; Pred. No. 6.2e-32;
Matches 120; Conservative 71; Mismatches 94; Indels 51; Gaps 9;

QY 1 MEKYEKLAKTGEISYGVVFKCRNKTSGQVAVKKF-VESEDDPVVKKIALREIRMLKQLK 59
DB 1 MEDYVKIEKIGEGTYGVVYKGRHKSTGQVAMKKIRLESEEG-VPSTAVREVSLQLQELK 59
QY 60 HPNLVNLIEVFRKRKMHVFEYCDHTLLNELERNP-NG--VADGVKSVLWQTLQALNFC 117
DB 60 HPNVVRLDDVLMQESRLYLIFEFLSMDLKKYLDIPSQGMDFMLVKSILYQILEGIYFC 119
QY 118 HIHNCIHRDIKPNENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDT 176
DB 120 HRRVLHRDLKPQNLLIDNKGVIKLADEGLARAFGVPRVTYTHEVVTWYRAPEVLLGSP 179
QY 177 QYGSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIRTLVETGFRHVDQAGLELLTSS 236
DB 180 RYSTPVDVWSTGTIFAELATKKPLFHGDSIDQLFRIFRTL-----G 221
QY 237 DP-----PAVASQSAGITGKLIPRHQSIFKSNQFFHGHSIPEPEDME-TLEEKFSVHPV 290
DB 222 TPNNDVWPDVES-----LPDYKNTF-----PKWMEGSLSSMVKNLDKN 259
```



```
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D17350; BAA04165.1; -.
DR HSSP; P24941; 1AQ1.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Cell cycle; Cell division; Mitosis; Phosphorylation.
FT DOMAIN 4 286 PROTEIN KINASE.
FT NP_BIND 10 18 ATP (BY SIMILARITY).
FT BINDING 33 33 ATP (BY SIMILARITY).
FT ACT_SITE 127 127 BY SIMILARITY.
FT MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 160 160 PHOSPHORYLATION (BY CAK) (BY SIMILARITY).
SQ SEQUENCE 298 AA; 33839 MW; D1LC036664C4C7CE CRC64;
Query Match 30.0%; Score 546; DB 1; Length 298;
Best Local Similarity 38.1%; Pred. No. 1.le-31;
Matches 117; Conservative 54; Mismatches 90; Indels 46; Gaps 6;
QY 1 MEKYEKLAKTGEISYGVVFKCRNKTSGQVAVKFESEDDFVVKKIALREIRMLKQLKH 60
DB 1 MENFQKVEKIGEGYGVVYKAKNKLGTGEVVALKKIRLDTETEGVPSTAIRISLLKELNH 60
QY 61 PNLVNLIEVFRKPKHVLVFEYCDHTLLNELERNP-NGVADGVIKSVLWQTLQALNFCHI 119
DB 61 PNIVKLLDVHTENKNLYLVFELLHQDLKKFMDASAVTGIPPLIKSYLFQLLQGLAFCHS 120
QY 120 HNCIHRDIKPNILITKGGIHKICDFGFAQL-IPGDAYTDYVATRWVRAPELLVGDTOY 178
DB 121 HRVLHRLKPKQNLINAEAGSIKLAQFLGLARAGVPVRTYTHEVVTWVRAPEILLGCKYY 180
QY 179 GSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTL----- 217
DB 181 STAVDIWSLGCIFAEWVTRRALFPDSEIDQLFRIFRTLTGTPDEVVWPGVTSMPDYKPSF 240
QY 218 ---VETGFRHV---DQAGLELLTSS---DPPAVASQSAGITGKLIPRHQSIFKSNQFFH 267
DB 241 PKWARQDFSKVPPLEDGSRLLSQMLHYDPNKRISAKAAL-----AHPFFQ 287
QY 268 GISIPEP 274
DB 288 DVTKPVP 294
RESULT 14
CDC2_HUMAN
IC CDC2_HUMAN STANDARD; PRT; 297 AA.
AC P06493; O60764;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Cell division control protein 2 homolog (EC 2.7.1.-) (p34 protein
kinase) (Cyclin-dependent kinase 2) (CDK1).
GN CDC2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87201915; PubMed=3553962;
RA Lee M.G.; Nurse P.;
RT "Complementation used to clone a human homologue of the fission yeast
```

```
RT cell cycle control gene cdc2.";
RL Nature 327:31-35(1987).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANT CDC2DELTAT.
RC TISSUE=Breast cancer;
RX MEDLINE=98175539; PubMed=9515786;
RA Ohta T., Okamoto K., Ischashi F., Shibata K., Fukuda M., Yamaguchi S.,
RA Xiong Y.;
RT "T-loop deletion of CDC2 from breast cancer tissues eliminates binding
to cyclin B1 and cyclin-dependent kinase inhibitor p21.";
RL Cancer Res. 58:1095-1098(1998).
RN [3]
RP SEQUENCE FROM N.A.
RA Rieder M.J., Livingston R.J., Braun A.C., Montoya M.A., Chung M.-W.,
RA Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,
RA Schackwitz W.S., Sherwood J.K., Witrak L.A., Nickerson D.A.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hu-lyk S.W.,
RA Villaion D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RP PHOSPHORYLATION, AND ASSOCIATION WITH P13.
RX MEDLINE=88253421; PubMed=3289755;
RA Draetta G., Beach D.;
RT "Activation of cdc2 protein kinase during mitosis in human cells:
cell cycle-dependent phosphorylation and subunit rearrangement.";
RL Cell 54:17-26(1988).
CC -!- FUNCTION: PLAYS A KEY ROLE IN THE CONTROL OF THE EUKARYOTIC
CELL CYCLE. IT IS REQUIRED IN HIGHER CELLS FOR ENTRY INTO S-PHASE
AND MITOSIS. P34 IS A COMPONENT OF THE KINASE COMPLEX THAT
PHOSPHORYLATES THE REPETITIVE CARBOXYL-TERMINUS OF RNA
POLYMERASE II.
CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
THE ENZYME, WHILE PHOSPHORYLATION AT THR-161 ACTIVATES IT.
CC -!- SUBUNIT: FORMS A STABLE BUT NON-COVALENT COMPLEX WITH A REGULATORY
SUBUNIT AND WITH A CYCLIN.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- POLYMERPHISM: Isoform CDC2delta is found in breast cancer
tissues. It is unable to complex with cyclin B1 and also fails to
bind to the CDK inhibitor p21.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CDC2/CDKX SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X05360; CAA28963.1; -.
```

DR EMBL; Y00272; CAA68375.1; -.
DR EMBL; D88357; BAA26001.1; -.
DR EMBL; AF512554; AAM34793.1; -.
DR EMBL; BC014563; AAH14563.1; -.
DR PIR; A29539; A29539.
DR PDB; 1LC9; 17-APR-02.
DR SWISS-2DPAGE; P06493; HUMAN.
DR Genew; HGNC:1722; CDC2.
DR GK; P06493; -.
DR MIM; 116940; -.
DR GO; GO:0005634; C:nucleus; TAS.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding; Cell cycle;
KW Cell division; Mitosis; Nuclear protein; Phosphorylation;
KW Polymorphism; 3D-structure.
FT DOMAIN 4 297 PROTEIN KINASE.
FT NP_BIND 10 18 ATP (BY SIMILARITY).
FT BINDING 33 33 ATP (BY SIMILARITY).
FT ACT_SITE 128 128 BY SIMILARITY.
FT MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 161 161 PHOSPHORYLATION (BY CAK) (BY SIMILARITY).
FT VARIANT 107 163 MISSING (IN CDC2delta).
FT /FTID=VAR_011629.
SQ SEQUENCE 297 AA; 34095 MW; 942D79448EFE490A CRC64;

Query Match 29.9%; Score 545; DB 1; Length 297;
Best Local Similarity 40.2%; Pred. No. 1.3e-31;
Matches 117; Conservative 58; Mismatches 74; Indels 42; Gaps 7;

QY 1 MEKYEKLAKTGECSYGWVFKCRNKTSGQVAVKKF-VESEDDPVVKKIALREIRMLKQLK 59
DB 1 MEDYTKIEKIGEGTYGVVYKGRHKTTGQVAVMKKIRLESEEG-VFSTAIRSILKELR 59

QY 60 HPNLVNLIEVFRKRKMHVFEYCDHTLNLNELERNFNG--VADGVKSVLWQTLQALNFC 117
DB HPNIVSLQVLMQDSRLYLIFEFLSMDLKKYLDSPPGQYMDSSAVKSYLYQLQGIVFC 119

QY 118 HHNCIHRDIPKENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGD 176
DB HRRVLHRDLKPQNLLDDKGTIKLADFGLARAFGPIRVYTHEVVTWYRSPEVLGSA 179

QY 177 QYGSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTL----- 217
DB RYSTPVDIWSIGTIFAELATKKPLFHGDSEIDQLFRIFRALGTPNNEVWPEVESJQDYKN 239

QY 218 -----VETGFRHVDQAGLELITSS--DPPAVASQSAGITCKLIPRH 256
DB TFPKWKPGSLASHVKNLDENGJDLISKMLIYDP-----AKRISGKMALNH 284

RESULT 15
CDK2_CARAU STANDARD; PRT; 298 AA.
AC P43450;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Cell division protein kinase 2 (EC 2.7.1.1-).
GN CDK2.
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Carassius.
OX NCBI_TaxID=7957;
RN [1]

RP SEQUENCE FROM N.A.
RC TISSUE=Oocyte;
RX MEDLINE=92331802; PubMed=1339336;
RA Hirai T., Yamashita M., Yoshikuni M., Tokumoto T., Kajiura H.,
RA Sakai N., Nagahama Y.,
RT "Isolation and characterization of goldfish cdk2, a cognate variant
of the cell cycle regulator cdc2.";
RL Dev. Biol. 152:113-120(1992).
CC -!- FUNCTION: PROBABLY INVOLVED IN THE CONTROL OF THE CELL CYCLE.
CC INTERACTS WITH CYCLINS A, D, OR E.
CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
THE ENZYME, WHILE PHOSPHORYLATION AT THR-160 ACTIVATES IT (BY
SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; S40289; AAB22550.1; -.
CC PIR; A44878; A44878.
CC HSSP; P24941; IFIN.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR002290; Ser_thr_pkinase.
CC Pfam; PF00069; pkinase; 1.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00220; S_TKC; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding; Cell cycle;
KW Cell division; Mitosis; Phosphorylation.
FT DOMAIN 4 296 PROTEIN KINASE.
FT NP_BIND 10 18 ATP (BY SIMILARITY).
FT BINDING 33 33 ATP (BY SIMILARITY).
FT ACT_SITE 127 127 BY SIMILARITY.
FT MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 160 160 PHOSPHORYLATION (BY CAK) (BY SIMILARITY).
SQ SEQUENCE 298 AA; 33998 MW; 50A93A63A235F9A9 CRC64;

Query Match 29.9%; Score 545; DB 1; Length 298;
Best Local Similarity 46.6%; Pred. No. 1.3e-31;
Matches 102; Conservative 48; Mismatches 67; Indels 2; Gaps 2;

QY 1 MEKYEKLAKTGECSYGWVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
DB 1 MESFQKVEKIGEGTYGVVYKAKNKVTGETVALKKIRLDTETEGVPSTAIRSILKELNH 60

QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLNLNELERNP-NGVADGVKSVLWQTLQALNFC 119
DB 61 PNIVKLHDVHTENKLYLVFEFLHQDLKRFMDSSVTGTGISLPLVKSYLFQLQLGLAFCHS 120

QY 120 HNCIHRDIPKENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGD 178
DB 121 HRVLHRDLKPQNLLINAQGEIKLADFGLARAFGVPVRTYTHEVVTWYRAPELLGCKY 180

QY 179 GSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTL 217
DB 181 STAVDIWSLGCIFAEMITRKALFPGDSEIDQLFRIFRTL 219

Search completed: November 13, 2003, 14:19:19
Job time : 11.4834 secs

GenCore version: 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 14:16:06 ; Search time 33.0227 Seconds
(without alignments)
2711.600 Million cell updates/sec

Title: US-09-671-050-10
Perfect score: 1820
Sequence: : MEKYEKLAKTGEGSYGVVFK.....RKARNEGRNRRRQQVLPKLS 347

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTRMBL_23:
1: sp_archea:
2: sp_bacteria:
3: sp_fungi:
4: sp_human:
5: sp_invertebrate:
6: sp_mammal:
7: sp_mhc:
8: sp_organelle:
9: sp_phage:
10: sp_plant:
11: sp_rodent:
12: sp_virus:
13: sp_vertebrate:
14: sp_unclassified:
15: sp_rvirus:
16: sp_bacteriap:
17: sp_archeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1180	64.8	352	11 Q8CEQ0	Q8ceq0 mus musculus
2	1001.5	55.0	353	5 Q9J12H1	Q9u2h1 caenorhabditis
3	1001.5	55.0	392	5 Q9VMN3	Q9vmn3 drosophila
4	941.5	51.7	566	6 Q9TTK0	Q9ttk0 cryptotlagus
5	915.5	50.3	493	4 Q92772	Q92772 homo sapien
6	904.5	49.7	564	11 Q9QY12	Q9qy12 mus musculus
7	904.5	49.7	568	11 Q9QUK0	Q9quk0 mus musculus
8	902	49.6	329	11 Q9QY11	Q9qy11 mus musculus
9	758.5	41.7	455	4 Q9P114	Q9p114 homo sapien
10	758.5	41.7	592	4 Q81VW4	Q81vw4 homo sapien
11	755	41.5	657	5 Q9BMG2	Q9bmg2 trypanosoma
12	744	40.9	353	11 Q8K134	Q8kl34 mus musculus
13	743.5	40.9	1106	5 Q9GRT9	Q9grt9 leishmania
14	741.5	40.7	457	11 Q8BKR2	Q8bkr2 mus musculus
15	738.5	40.6	585	11 Q8BL49	Q8bl49 mus musculus
16	737.5	40.5	457	11 Q9Jm02	Q9jm02 rattus norv

17	737.5	40.5	505	11 Q9JmC1	Q9jm01 rattus norv
18	733.5	40.3	595	11 Q8BLF2	Q8blf2 mus musculu
19	694.5	38.2	783	11 Q8BWI8	Q8bwi8 mus musculu
20	694.5	38.2	831	4 Q8IYC7	Q8iyc7 homo sapien
21	682	37.5	993	4 Q8WXQ5	Q8wxq5 homo sapien
22	676.5	37.2	578	11 Q8BVE0	Q8bve0 mus musculu
23	618	34.0	1104	13 Q9W6R6	Q9w6r6 fugu rubrip
24	615.5	33.8	997	4 Q9UJL6	Q9ujl6 homo sapien
25	565.5	31.1	294	10 Q9ZRI1	Q9zrl1 triticum ae
26	551	30.3	288	5 Q96821	Q96821 plasmodium
27	552	30.3	288	5 Q9XZD6	Q9xzd6 plasmodium
28	545.5	30.0	294	10 Q82666	Q82666 brassica na
29	542.5	29.8	300	5 Q15890	Q15890 toxoplasma
30	539.5	29.6	294	10 Q40790	Q40790 pinus conto
31	539.5	29.6	297	10 Q8RX68	Q8rx68 arabidopsis
32	539.5	29.6	300	5 Q17066	Q17066 asterina pe
33	539.5	29.6	300	5 Q44000	Q44000 toxoplasma
34	539	29.6	298	5 Q27032	Q27032 theileria p
35	538.5	29.6	294	10 Q40789	Q40789 petroselinu
36	538.5	29.6	294	10 Q82135	Q82135 pisum sativ
37	538.5	29.6	300	3 Q13379	Q13379 pneumocysti
38	538.5	29.6	300	3 Q13380	Q13380 pneumocysti
39	537	29.5	288	5 Q96820	Q96820 plasmodium
40	535.5	29.4	294	10 Q43361	Q43361 picea abies
41	535.5	29.4	294	10 P93556	P93556 sesbania ro
42	534.5	29.4	294	10 Q9FUR4	Q9fur4 nicotiana t
43	534.5	29.4	294	10 Q40484	Q40484 nicotiana t
44	534	29.3	298	5 Q26671	Q26671 theileria a
45	533.5	29.3	294	10 Q8L6T8	Q8l6t8 daucus caro

ALIGNMENTS

RESULT 1
Q8CEQ0
ID Q8CEQ0 PRELIMINARY; PRT; 352 AA.
AC Q8CEQ0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Cyclin-dependent kinase-like 1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK016781; BAC25497.1; -.
SQ SEQUENCE 352 AA; 41023 MW; 6E9F56C8080F35E0 CRC64;

QY	1	MEKYEKLAKTGEGSYGVVFKRNKTSQVAVVKFVSEDDPVVKKIALREIRMLKQLKH	60
Db	1	MEKYEKIGKIGEGSYGVVFKRNRTGQIVAIKRFLETEDDPVIKKIALREIRMLKQLKH	60
QY	61	PNLVNLIEVFRKRKRMHLVFEYCDHTLLNELEPNPNVADGVKSVLWQTLQALNFCIH	120
Db	61	PNLVNLIEVFRKRRLRLHLVFEYCDHTLVHELDRYQRGVPEPLVKNTWTQTLQAVNFKKH	120
QY	121	NCIHRDIKPENILITKQGIKICDFGFAQILI-PGDAYTDYVATRWYRAPELLVGDYQG	179
Db	121	NCIHRDVKPENILITKQSAIKLCDFGARLLTGPGDYYTDYVATRWYRSPELLVGDYQG	180

Query Match 64.8%; Score 1180; DB 11; Length 352;
Best Local Similarity 62.7%; Pred. No. 6.1e-98;
Matches 222; Conservative 42; Mismatches 48; Indels 42; Gaps 4;

```
QY 180 SSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIPTLVETGFRHVDQAGLELLTSSFP 239
Db 181 PVDVWAIGCVFAELLGVPWPGKSDVDQLYLIRKTL----- 219
QY 240 AVASQSAGITGKLI PRHQSIFKSNQGFHGISIPEPEDETLEEKFSVDHPVALNFMKGLK 299
Db 219 -----GDLIPRHQQVFSMNQYFSGVKIPDPEDMETLELKFNPISYALGELKGC 268
QY 300 KMPDRLTCSQLLESSYFDSFQEA-----QKRRKARNEGRRRQ-----QVLP 344
Db 269 HMDPAERLTCEQLLQHPYFDSIREVGELTRQHXDKPARKTLRQSRKHLTG-QYLP 322

RESULT 2
Q9U2H1
ID Q9U2H1; PRELIMINARY; PRT; 353 AA.
AC Q9U2H1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Y42A5A.4 protein.
GN Y42A5A.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Steward C.A.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RM [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851915;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AL032618; CAB63367.1; -.
DR HSSP; P24941; IHCL.
DR WormPep; Y42A5A.4; CE20258.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD00C001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 353 AA; 40665 MW; 7E3226D09849CAC9 CRC64;

Query Match 55.0%; Score 1001.5; DB 5; Length 353;
Best Local Similarity 53.4%; Pred. No. 7.5e-82;
Matches 187; Conservative 59; Mismatches 63; Indels 41; Gaps 3;

QY 1 MEKYEKAKTGEQSYGVWFKCRNKTSGQWAVKKEVEDDPVYKIKALREIRMLKQLKH 60
Db 1 MDKYDRLSKLGEQSYGVWYKCNKRDGTQIVAIKKFVETEDPHIKKIALREIRMLKQLKH 60
QY 61 PNLVNLIEVFRKRMHLVFECYCDHTLLNELERNPNQVADGVKISVLWQTLQALNFCYH 120
Db 61 QNLVGLIEVFRKRMHLVFLCDRTVLHLEKXPNHGVNDELIIKKIYQLLEALKFCHSH 120
QY 121 NCIHRIKIPENILITKQIIKICDFGFAQLIPGDYATDYVATRWYRAPPELLVGDYQYS 180
Db 121 KCIHRDVKPENIFLTRNDQVKLGDFGFARIINTTEMTDYVATRWYRSPPELLVGDYQYP 180
QY 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIPTLVETGFRHVDQAGLELLTSSOPPA 240
Db 181 PVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIPTLVETGFRHVDQAGLELLTSSOPPA 217
```

```
QY 241 VASQSAGITGKLI PRHQSIFKSNQGFHGISIPEPEDETLEEKFSVDHPVALNFMKGLK 300
Db 218 -----GEFLPRHISIFRTNQFFGLSIPEPEHLEPLSKLPNASSAQDLFLQKCFE 268
QY 301 MNPDDRILTCSQLLESSYFDSFQEAQIKRKARNEG-----RNRRRQQVLP 345
Db 269 MSPDRRFSCSELMLHGIFSNW-----ILRIQDESTPTGLTSKRSPNYLPL 314

RESULT 3
Q9VMN3
ID Q9VMN3; PRELIMINARY; PRT; 392 AA.
AC Q9VMN3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE CG7236 protein.
GN CG7236.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=107311132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA War K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadiou E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.C., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AE003611; AAF52279.1; -.
DR HSSP; P24941; IHCL.
DR FlyBase; FBgn0031730; CG7236.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
```



```
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 392 AA; 44962 MW; C3627D077784E55D CRC64;

Query Match 55.0%; Score 1001.5; DB 5; Length 392;
Best Local Similarity 53.0%; Pred. No. 8.7e-82;
Matches 184; Conservative 65; Mismatches 61; Indels 37; Gaps 3;

QY 1 MEKYEKLAKTGECSYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MDRYEKLRLGEGSYGVVYKCRDRTGALVAVKRFVESEDDPAIRKIALREIRMLKQLKH 60

QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEARNPNGVADGVIKSVLWQTLQALNFCCHI 120
Db 61 PNLVSLIEVFRKRRLHLVFEFCETVLHELERHPOGCPHEHTKQICYQTLGVAICHKQ 120

QY 121 NCIHRDIKPENILITKQGIKICDFGFAQILIPGDAYTDYVATRWYRAPELLVGDTOYG 150
Db 121 GCLHRDIKPENILITAQGVKLCDFGFARMLSPGENYTDYVATRWYRAPELLVGDTOYG 150

QY 181 SVDIWAIGCVFAELLTGQPLWPKSDVDQLYLIRTLVETGFRHVDQAGJELLTSSDPPA 240
Db 181 PVDVWAIGCLFAELVRGEALWPKRSDVDQLYLIRKTL----- 217

QY 241 VASQSAGITGKLIPRHQSIFKSNGFPHGISIPEPEDMETLEEKF---SDVHPVALNFMKG 297
Db 218 -----GDLLPRHIQIFGQNEVFKGITLPVPPTLEPLECKVPAKSQQNPLTIDFLKK 268

QY 298 CLGKNPDDLTCSSQLLESSYFDSF--QEAQIKRKARNEGRRRRQQV 342
Db 269 CLDKDPTKWSCEKLTKHSYFDDYIAKQRELEHNSLEAANLRQQQL 315

RESULT 4
Q9TTK0 PRELIMINARY; PRT; 566 AA.
AC Q9TTK0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Ser/Thr kinase KKIAMRE.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Gomi H., Sun W., Finch C.E., Itohara S., Yoshimi K., Thompson R.F.;
RT "Learning induces a cdc2-related protein kinase, kkiamre.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB029045; BAA88508.1; -.
DR HSSP; P24941; IHCL.
DR InterPro; IPR000719; Prot_kinase.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 566 AA; 64052 MW; 5D57595550902EA9 CRC64;

Query Match 51.7%; Score 941.5; DB 6; Length 566;
Best Local Similarity 49.7%; Pred. No. 3.7e-76;
Matches 177; Conservative 67; Mismatches 65; Indels 47; Gaps 6;

QY 1 MEKYEKLAKTGECSYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
```

```
Db 1 MEKYENLGLVGECSYGMVKCRNKDSGRIVAIAKFLESDDDDKMKVKKIAMREIKLLKQLRH 60
QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEARNPNGVADGVIKSVLWQTLQALNFCCHI 120
Db 61 ENJVNLLJEVCKKRWLVFEFVDHTILDDLELFPNGLDDQVVQVLYFQIINGIGFCHSH 120

QY 121 NCIHRDIKPENILITKQGIKICDFGFAQILIPGDAYTDYVATRWYRAPELLVGDTOYG 179
Db 121 NCIHRDIKPENILVSQSGVVKLCDFGFARTLAAPGEVYTDYVATRWYRAPELLVGDVKYG 180

QY 180 SVDIWAIGCVFAELLTGQPLWPKSDVDQLYLIRTLVETGFRHVDQAGLELLTSSDPP 239
Db 181 KAVDWAIGCLVTEMVMGEPLFPGDSIDQLYLIMRCL----- 218

QY 240 AVASQSAGITGKLIPRHQSIFKSNGFPHGISIPEPEDMETLEEKFSVDVHPVALNFMKG 299
Db 219 -----GNLIPRHELFYKNPVFAGVRLPEIKSEPLERRYPKLSEVVIDLAKKCL 268

QY 300 KKNFDDRLTCSQLLESSYF--DSF-----QEAQIK--RKARN-----EGRNRRRQQ 341
Db 269 HVDPDKRPFCALLHRDFFQMDGFAERFSQELQMKVQKDARNISLSKKSQNRKKEK 324

RESULT 5
Q92772 PRELIMINARY; PRT; 493 AA.
AC Q92772;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE P56 KKIAMRE protein kinase.
GN KKIAMRE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97152547; PubMed=9000130;
RA Taglienti C.A., Wysk M., Davis R.J.;
RT "Molecular cloning of the epidermal growth factor-stimulated protein
RT kinase p56 KKIAMRE.";
RL Oncogene 13:2563-2574(1996).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; U35146; AAC50918.1; -.
DR HSSP; P24941; IHCL.
DR Genew; HGNC:1782; CDKL2.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 493 AA; 56018 MW; 5CC20A91CBF89EFE CRC64;

Query Match 50.3%; Score 915.5; DB 4; Length 493;
Best Local Similarity 48.3%; Pred. No. 6.8e-74;
Matches 172; Conservative 68; Mismatches 69; Indels 47; Gaps 6;

QY 1 MEKYEKLAKTGECSYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MEKYENLGLVGECSYGMVKCRNKDTRIVAIKAFLESDDDDKMKVKKIAMREIKLLKQLRH 60

QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEARNPNGVADGVIKSVLWQTLQALNFCCHI 120
Db 61 ENLVNLLJEVCKKRWLVFEFVDHTILDDLELFPNGLDYQVVQVLYFQIINGIGFCHSH 120

QY 121 NCIHRDIKPENILITKQGIKICDFGFAQILIPGDAYTDYVATRWYRAPELLVGDTOYG 179
Db 121 NCIHRDIKPENILITKQGIKICDFGFAQILIPGDAYTDYVATRWYRAPELLVGDTOYG 179
```

Db 121 NIIHRDIKPENILVSQGVVKLCDFGFARTLAAPGEVYTDYVATRWYRAPELLVGVKYG 180
QY 180 SSVDIWAIGCVFAELLTQGPLWPGKSDVDQLYJIRTLVETGFRHVDQAGJELLTSSDPP 239
Db 181 KAVDVWAIGCLVTEMFMGEPLFPGDSIDIDQJYHMMCL----- 218
QY 240 AVASQSAGITKGLIPRHQSIFKSNNGFFHGISIPEDMETLEEKFSVHPVVALNFMKGC 299
Db 219 -----GNLIPRHQELFNKNPVFAGVRLPEIKEREPLERRYPKLSEVVIDLAKKCL 268
QY 300 KMNPPDRRLTCSQLLESSYF--DSF-----QEAQIK--RKARN-----EGRNRRRQQ 341
Db 269 HIDPKRPFCAELLHHDDFFQMDGFAERFSQELQLKVQKQARNVSLSKXSONRKKEK 324
RESULT 6
Q9QYI2 PRELIMINARY; PRT; 564 AA.
AC Q9QYI2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Ser/Thr kinase KKIAMRE-gamma.
GN CDKL2 OR KKY.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Sassa T., Gomi H., Sun W., Ikeda T., Thompson R.F., Itohara S.;
RT "The Murine KKIAMRE gene: variants, dual promoters, expression and
RT chromosomal localization."
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB029066; BAA88428.1; -.
DR HSSP; P24941; IHCL.
DR WGD; MGI:1858227; Cdk12.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PFC0069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 564 AA; 63640 MW; DACDEF630CCE82D5 CRC64;

Query Match 49.7%; Score 904.5; DB 11; Length 564;
Best Local Similarity 48.6%; Pred. No. 8.le-73;
Matches 174; Conservative 56; Mismatches 69; Indels 49; Gaps 7;
QY 1 MEKYEKLAKTGEYSYGVVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MEKYENLGLVGEYSYGMVKCRNKDSGRIVA-KKFLSEDDDKVVKKIARREIKLLKQLRH 60
QY 61 PNLVNLIEVFRRRKRMHLVFEYCDHTLLNELERNPNGVADGVIKSVLWQTLOALNFCIH 120
Db 61 ENLVNLLEVCKKKRWLVFEFVDHTILDCLKLPNGLDYQVWQKYLFIQINGIGFCHSH 120
QY 121 NCIHRDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDQY 179
Db 121 NIIHRDIKPENILVSQGVVKLCDFGFARTLAAPGEVYTDYVATRWYRAPELLVGDVKG 180
QY 180 SSVDIWAIGCVFAELLTQGPLWPGKSDVDQLYJIRTLVETGFRHVDQAGJELLTSSDPP 239
Db 181 KAVDIWAIGCLVIELMGLQPLFPGESDIDQLHHIMTCL----- 218
QY 240 AVASQSAGITKGLIPRHQSIFKSNNGFFHGISIPEPEDMET--LEEKFSVHPVVALNFMKG 297
Db 219 -----GNLIPRHQELFYKNPVFAGVRLPEVKDAEAEPLERSYPKLPEAVISLAKK 268

QY 298 CLKMNPPDRRLTCSQLLESSYF--DSF-----QEAQIK--RKARN-----EGRNRRRQQ 341
Db 269 CLHIDPKRPFCAELLRHDDFFQMDGFAERFSQELQLKIEKDARNNSLPKKSQNRKKEK 326
RESULT 7
Q9QUK0 PRELIMINARY; PRT; 568 AA.
AC Q9QUK0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Ser/Thr kinase KKIAMRE.
GN CDKL2 OR KKY.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ, and C57BL/6;
RA Sassa T., Gomi H., Sun W., Ikeda T., Thompson R.F., Itohara S.;
RT "The Murine KKIAMRE gene: variants, dual promoters, expression and
RT chromosomal localization."
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB029073; BAA88439.1; -.
DR EMBL; AB029068; BAA88439.1; JOINED.
DR EMBL; AB029069; BAA88439.1; JOINED.
DR EMBL; AB029070; BAA88439.1; JOINED.
DR EMBL; AB029071; BAA88439.1; JOINED.
DR EMBL; AB029072; BAA88439.1; JOINED.
DR EMBL; AB029065; BAA88427.1; -.
DR HSSP; P24941; IHCL.
DR XGD; MGI:1858227; Cdk12.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PFC0069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 568 AA; 64055 MW; A43B75E2E9EB86C4 CRC64;

Query Match 49.7%; Score 904.5; DB 11; Length 568;
Best Local Similarity 48.6%; Pred. No. 8.le-73;
Matches 174; Conservative 66; Mismatches 69; Indels 49; Gaps 7;
QY 1 MEKYEKLAKTGEYSYGVVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MEKYENLGLVGEYSYGMVKCRNKDSGRIVA-KKFLSEDDDKVVKKIARREIKLLKQLRH 60
QY 61 PNLVNLIEVFRRRKRMHLVFEYCDHTLLNELERNPNGVADGVIKSVLWQTLOALNFCIH 120
Db 61 ENLVNLLEVCKKKRWLVFEFVDHTILDCLKLPNGLDYQVWQKYLFIQINGIGFCHSH 120
QY 121 NCIHRDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDQY 179
Db 121 NIIHRDIKPENILVSQGVVKLCDFGFARTLAAPGEVYTDYVATRWYRAPELLVGDVKG 180
QY 180 SSVDIWAIGCVFAELLTQGPLWPGKSDVDQLYJIRTLVETGFRHVDQAGLELLTSSDPP 239
Db 181 KAVDIWAIGCLVIELMGLQPLFPGESDIDQLHHIMTCL----- 218
QY 240 AVASQSAGITKGLIPRHQSIFKSNNGFFHGISIPEPEDMET--LEEKFSVHPVVALNFMKG 297
Db 219 -----GNLIPRHQELFYKNPVFAGVRLPEVKDAEAEPLERSYPKLPEAVISLAKK 268
QY 298 CLKMNPPDRRLTCSQLLESSYF--DSF-----QEAQIK--RKARN-----EGRNRRRQQ 341

Db 269 CLHIDPKRPF CADLLRHDFQMDGFAERFSQELQLKIEKDARNNSLPKKSQNRKKEK 326

RESULT 8

Q9QYI1 PRELIMINARY; PRT; 329 AA.

AC Q9QYI1;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Ser/Thr kinase KIAMRE-delta (fragment).

GN CDKL2 OR KKM.

CS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6; TISSUE=Brain;

RA Sassa T., Gomi H., Sun W., Ikeda T., Thompson R.F., Itohara S.;

RT "The Murine KIAMRE gene: variants, dual promoters, expression and

RT chromosomal localization."

RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.

CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

DR EMBL; AB029067; BAA88429.1; -.

DR HSSP; P24941; IHCL.

DR MGD; MGI:1858227; Cdkl2.

DR InterPro; IPR000719; Prot_kinase.

DR InterPro; IPR002290; Ser_thr_pkinase.

DR Pfam; PF00069; pkinase; 1.

DR ProDom; PD000001; Prot_kinase; 1.

DR SMART; SM00220; S_TKC; 1.

DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.

KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.

FT NON TER 329

SQ SEQUENCE 329 AA; 38023 MW; F72EB89E65AE8585 CRC64;

Query Match 49.6%; Score 902; DB 11; Length 329;

Best Local Similarity 49.0%; Pred. No. 6.4e-73;

Matches 173; Conservative 63; Mismatches 73; Indels 44; Gaps 6;

QY 1 MEKYEKLAKTGE GSYGVVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIRMLKQLKH 60

Db 1 MEKYENLGLVGE GSYGMVMKCRNKDSGRIVAIKFLFESDDDRKVKIAMREIKLLQLRH 60

QY 61 PNLVNLIEVFRKRKXKHLVFEYCDHTLLNELERNPNGVADGVKSVLWQTLQALNFCIH 120

Db 61 ENLVNLLEVCCKKRWYLVFEFVDHTILDLLKLPNGLDYQVQVKYLEFQIINGIGFCHSH 120

QY 121 NCIHRLDKPENILITKQGIKICDFGPAQL-IPGDAYTDYVATRWYRAPELLVGDYQY 179

Db 121 NIHRDIKPENILVVSQSGVVKLCDFGFARTLAAPGEVTDYVATRWYRAPELLVGDVKY 180

QY 180 SSVDIWAIGCVFAELLTGOPLPWPKSDVDQLYLIIRCTLVETGFRHVDQAGLELLTSSDPP 239

Db 181 KAVDIWAIGCLVIEMLMGQPLFPGESIDQLHIMTCL----- 218

QY 240 AVASQSAGITGKLI PRHQSIKSNGFHGISIPEPEDMET--LEEKFSVHPVALNFMKG 297

Db 219 -----GNLIPRHQELFYKNPVFAGVRLPEVKCAEAEPLESRYPKLPEAVSLAKK 268

QY 298 CLKMNPDDRLTCSQLESSEYF--DSF-----CEAQIK--RKARNEGRNRRQ 341

Db 269 CLHIDPKRPF CADLLRHDFQMDGFAERFSQELQLKIEKDARNNSLPKKSQK 321

RESULT 9

Q9P114 PRELIMINARY; PRT; 455 AA.

AC Q9P114;

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Serine-threonine protein kinase NKIAMRE.

GN NKIAMRE.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Midmer M., Hag R., Zanke B.W.;

RT "NKIAMRE a novel kinase deleted in human leukemia."

RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.

CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

DR EMBL; AF130372; AAF36509.1; -.

DR HSSP; P24941; IHCL.

DR Genew; HGNC:15483; CDKL3.

DR InterPro; IPR000719; Prot_kinase.

DR InterPro; IPR002290; Ser_thr_pkinase.

DR Pfam; PF00069; pkinase; 1.

DR ProDom; PD000001; Prot_kinase; 1.

DR SMART; SM00220; S_TKC; 1.

DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.

KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.

SQ SEQUENCE 455 AA; 51566 MW; 12046F84A363F7A0 CRC64;

Query Match 41.7%; Score 758.5; DB 4; Length 455;

Best Local Similarity 44.0%; Pred. No. 8.7e-60;

Matches 150; Conservative 61; Mismatches 91; Indels 39; Gaps 5;

QY 1 MEKYEKLAKTGE GSYGVVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIRMLKQLKH 60

Db 1 MEMYETLGKVGEGSYGTVMKCKHKTGTQIVAIKIFVERPEQS-VNKIAMREIKFLKQFHH 59

QY 61 PNLVNLIEVFRKRKXKHLVFEYCDHTLLNELERNPNGVADGVKSVLWQTLQALNFCIH 120

Db 60 ENLVNLIEVFRQKKIHLVFEFIDHTVLDLQHYCHGLESKRLRKYLEQLRAIDYLSHN 119

QY 121 NCIHRLDKPENILITKQGIKICDFGPAQL-IPGDAYTDYVATRWYRAPELLVGDYQY 179

Db 120 NIHRDIKPENILVVSQSGITKLCDFGFARTLAAPGDIYDQVATRWYRAPELLVKDTSY 179

QY 180 SSVDIWAIGCVFAELLTGOPLPWPKSDVDQLYLIIRCTLVETGFRHVDQAGLELLTSSDPP 239

Db 180 KPVDIWAIGCVFAELLTGOPLPWPKSDVDQLYLIIRCTLVETGFRHVDQAGLELLTSSDPP 217

QY 240 AVASQSAGITGKLI PRHQSIKSNGFHGISIPEPEDMETLEEKFSVHPVALNFMKG 299

Db 218 -----GNLSPHLQNIFFSKSPIFAGVVLPGVQHPKNARKKYPKLGKLGADIVHACL 267

QY 300 KKNPDRLTCSQLESSEYF--DSFQE---AQIKRKARNEGR 335

Db 268 QIDPADRISSDDLHHEHYFTRDGFIEKFMPELAKLLQEA 308

RESULT 10

Q8IWM4 PRELIMINARY; PRT; 592 AA.

ID Q8IWM4;

AC Q8IWM4;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Similar to cyclin-dependent kinase-like 3.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Testis;

RA Strausberg R.;

Db 180 KPVDIWALGCMYIEMATGHPFLPSSDDLHLHKIVLKV----- 217

QY 240 AVASQASAGITGKLIIPRHQSIFKSNGFHGISIPEPECVETLEKFSVHPVALNFMKGCL 299

Db 218 -----GNLTPHLNIFSKSPIFAGVLPQVQHPTARKKYPKLNGLLADIVHACL 267

QY 300 KMNPDRLTCSQLLESSYF--DSFQEAQ-----IKRKARNEGRNRRCQ 342

Db 268 QIDPAERTSSTDLLRHDYFTRDGFIEKKWTEKGPSSKSESRLPKGAKEMSOTRSQ 325

RESULT 13

Q3GRT9 PRELIMINARY; PRT; 1106 AA.

AC Q3GRT9;

DT 01-MAR-2001 (TrEMBLrel. 16, Created)

DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Putative mitogen-activated protein kinase 6.

GN MPK6.

OS Leishmania mexicana.

OC Eukaryota; Euzoenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.

OX NCBI_TaxID=5665;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MNYC/BZ/62/M379;

RA Wiese M., Wang Q., Goercke I.;

RT "Identification of mitogen-activated protein (MAP) kinases from Leishmania mexicana."

RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.

CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

DR EMBL; AJ293284; CAC07960.1; ..

DR HSSP; P24941; iB38.

DR InterPro; IPR000719; Prot_kinase.

DR InterPro; IPR002290; Ser_thr_pkinase.

DR Pfam; PF00069; pkinase; i.

DR ProDom; PD000001; prot_kinase; i.

DR SMART; SM00220; S_TKC; 1.

DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.

KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.

SQ SEQUENCE 1106 AA; 118884 MW; 24AC26F5C209923E CRC64;

Query Match 40.9%; Score 743.5; DB 5; Length 1106;

Best Local Similarity 43.7%; Pred. No. 6.8e-58;

Matches 145; Conservative 62; Mismatches 88; Indels 37; Gaps 3;

QY 1 MEKYEKLAKTGECSYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60

Db 1 MEAYETLGILGEGTYGVVVKARSRTGKLVAKRFKQTEQDEHVRKTSRVRMLQLLOH 60

QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELERNP2NGVADGVKSVLWQTLQALNFCIH 120

Db 61 PNVRLEDVFRREGKLYLVFEFIDHTILQLLESTTRGFRHRELRRYTYQLLRGIEFCHNQ 120

QY 121 NC1HRDIKPENILITKQGIKICDFGFAQILIPGDAYTDYVATRWYRAPELLVGDYQYS 180

Db 121 NIHRDVKPNVNLIDESGLKLCDFGFARQTSAGKGYTDYVATRWYRAPELLVGDVAYGK 180

QY 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLVETGFRHVDQAGLELLTSSDPPA 240

Db 181 PVDVWALGCMFAELSDGQPLFGESDLDQLCLIMQT----- 216

QY 241 VASQASAGITGKLIIPRHQSIFKSNGFHGISIPEPECVETLEKFSVHPVALNFMKGCLK 300

Db 217 -----CGVPVQRLVFIEMHNPPLYNGISFPHTDILYTLKDRYHRESNDWTEFLSSCLH 268

QY 301 MNPDDRLTCSQLLESSYF--DSFQ---EAQIK 327

Db 269 TDPAQLTCTELMELPYFTRDGRDRYEALR 300

RESULT 14

Q8BKR2 PRELIMINARY; PRT; 457 AA.

AC Q8BKR2;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Serine/threonine kinase NKIATRE alpha homolog.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Body;

RX MEDLINE=22354683; PubMed=12466851;

RA The FANTOM Consortium,

RT the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";

R1 Nature 420:563-573 (2002).

DR EMBL; AK050950; BAC34498.1; ..

SQ SEQUENCE 457 AA; 51900 MW; 6FCDD36F51152B61 CRC64;

Query Match 40.7%; Score 741.5; DB 11; Length 457;

Best Local Similarity 42.8%; Pred. No. 3e-58;

Matches 146; Conservative 64; Mismatches 92; Indels 39; Gaps 5;

QY 1 MEKYEKLAKTGECSYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60

Db 1 MEMYETLGKVGEGSYGTVMCKKHDTGRIVAIKIPYE-KPEKSVNKIATREIKFLKQFRH 59

QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELERNP2NGVADGVKSVLWQTLQALNFCIH 120

Db 60 ENLVNLIEVFRQKKIHLVFEFIDHTLVDELQHYCHGLESKRLRKYLQILRAISYLHN 119

QY 121 NC1HRDIKPENILITKQGIKICDFGFAQILIPGDAYTDYVATRWYRAPELLVGDYQYG 179

Db 120 NIHRDIKPENILVSQSGITKLCDFGFARTLAAPGDVYTDYVATRWYRAPELLVKTYSY 179

QY 180 SSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLVETGFRHVDQAGLELLTSSDPP 239

Db 180 KPVDIWALGCMYIEMATGHPFLPSSDDLHLHKIVLKV----- 217

QY 240 AVASQASAGITGKLIIPRHQSIFKSNGFHGISIPEPECVETLEKFSVHPVALNFMKGCL 299

Db 218 -----GNLTPHLNIFSKSPIFAGVLPQVQHPTARKKYPKLNGLLADIVHACL 267

QY 300 KMNPDRLTCSQLLESSYF--DSFQ---AQIKRKARNEGR 335

Db 268 QIDPAERTSSTDLLRHDYFTRDGFIEKFIPELRAKLQLEAK 308

RESULT 15

Q8BL49 PRELIMINARY; PRT; 585 AA.

AC Q8BL49;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Serine/threonine kinase NKIATRE alpha homolog.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Brain;

RX MEDLINE=22354683; PubMed=12466851;

RA The FANTOM Consortium,

RT the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM: protein - protein search, using sw model

Run on: November 13, 2003, 14:12:21 ; Search time 36.1631 Seconds
(without alignments)
1382.592 Million cell updates/sec

Title: US-09-671-050-12
Perfect score: 1662
Sequence: 1 MEKYEKLAKTGSGSYGVFK.....RKARNEGRNRRQVJPLKS 315

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03.*
1: /SIDSl/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSl/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSl/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSl/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSl/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSl/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSl/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSl/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSl/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSl/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSl/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSl/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSl/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSl/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSl/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSl/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSl/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSl/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSl/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSl/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSl/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSl/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSl/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSl/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1662	100.0	315	AAE00495	Human kinase #6.
2	1639	98.6	342	AAE19152	Human kinase polyp
3	1638	98.6	324	AAE00492	Human kinase #3.
4	1636	98.4	347	AAE00494	Human kinase #5.
5	1619	97.4	360	AAJ03525	Human protein: kina
6	1612	97.0	356	AAE00491	Human kinase #2.
7	1296.5	78.0	296	AAE65643	Novel protein kina
8	1188.5	71.5	358	ASP62954	Human polypeptide
9	1149	69.1	247	AAE65642	Novel protein: kina

10	1140	68.6	228	23	AAG78547	Human kinase 14257
11	1027.5	61.8	392	22	ABB63118	Drosophila melanog
12	967.5	58.2	566	21	AAV90724	Rabbit KIAMRE kin
13	948	57.0	197	24	ABU11689	Human MDDT polypep
14	937.5	56.4	362	24	ABP96087	Human protein kina
15	791	47.6	187	22	AAE00490	Human kinase #1.
16	791	47.6	198	22	AAE00493	Human kinase #4.
17	784.5	47.2	455	21	AAV70126	Human lost in leuk
18	784.5	47.2	591	22	AAU03524	Human protein kina
19	763.5	45.9	457	21	AAV70125	Rat lost in leukae
20	763.5	45.9	505	21	AAV70124	Rat lost in leukae
21	712.5	42.9	154	24	ABU05290	Human diagnostics
22	656	39.5	534	22	AAE65645	Novel protein kina
23	581	35.0	333	22	AAE39276	Human polypeptide
24	574	34.5	352	22	AAE41062	Human polypeptide
25	570	34.3	298	20	AAW95689	Homo sapiens Cdc2
26	570	34.3	544	19	AAW85028	CDK2-green fluore
27	570	34.3	544	19	AAW85029	Green fluorescent
28	569	34.2	224	22	ABG06142	Novel human diagno
29	568.5	34.2	294	20	AAW95690	Oryza sativa Cdc2
30	568.5	34.2	297	23	ABE79004	Human 10Ckshs1 pro
31	566.5	34.1	294	21	AAG33356	Zea mays protein f
32	565.5	34.0	294	21	AAG40809	Zea mays protein f
33	561.5	33.8	294	21	AAG11214	Arabidopsis thalia
34	561.5	33.8	294	21	AAG54042	Arabidopsis thalia
35	561.5	33.8	350	21	AAG11213	Arabidopsis thalia
36	557	33.5	297	20	AAW95697	Rattus norvegicus
37	556	33.5	297	20	AAW95688	Homo sapiens Cdc2
38	556	33.5	297	22	AAG67434	Amino acid sequenc
39	556	33.5	297	22	AAG67613	Amino acid sequenc
40	556	33.5	297	24	ABR47405	Breast cancer asso
41	556	33.5	297	24	AAE34491	Human CDC2 protein
42	556	33.5	297	24	ABU56522	Lung cancer-associ
43	555.5	33.4	300	20	AAW95679	Pneumocystis carin
44	555	33.4	270	20	AAV43932	Human protein kina
45	554	33.3	298	21	AAV52184	Cyclin-dependent k

ALIGNMENTS

RESULT 1:
AAE00495
ID AAE00495 standard; Protein; 315 AA.
XX
AC AAE00495;
XX
DT 19-JUN-2001 (first entry)
XX
DE Human kinase #6.
XX
KW Human; kinase; gene therapy; bioreactor; mental disorder;
KW biological disorder.
XX
OS Homo sapiens.
XX
PN WQ200123579-A1.
XX
PD 05-APR-2001.
XX
PF 27-SEP-2000; 2000WO-US26621.
PR 28-SEP-1999; 99US-0156511.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
XX
DR WPI; 2001-266166/27.
DR N-PSDB; AAD03817.
XX
PT New isolated human kinase polynucleotide useful for generating
PT antibodies, as reagents in diagnostic assays and for screening for

Db 1 MEKYEKLAKTGEYSYGVVFKCRNKTSGQVVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEPNPVGADGVKSVLWQTLQALNFCIH 120
Db 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEPNPVGADGVKSVLWQTLQALNFCIH 120
QY 121 NCIHRIKIPENILITKQGIKICDFGFAQILIPGDAYDYVATRWYRAPELLVGDTCYGS 180
Db 121 NCIHRIKIPENILITKQGIKICDFGFAQILIPGDAYDYVATRWYRAPELLVGDTCYGS 180
QY 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIIPRHQSIFKSNGFHGISIP 240
Db 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIIPRHQSIFKSNGFHGISIP 240
QY 241 EPEDMETLEEKFSVDVHPVALNFMKGLKMPDRLTCSQLLESSYFDSFQEAQIKRKARN 300
Db 241 EPEDMETLEEKFSVDVHPVALNFMKGLKMPDRLTCSQLLESSYFDSFQEAQIKRKARN 300
QY 301 EGRNRRRQ--QVLP 313
Db 301 EGRNRRRQ--QVLP 313

RESULT 3
AAE00492
ID AAE00492 standard; Protein; 324 AA.

AC AAE00492;
XX 19-JUN-2001 (first entry)
XX Human kinase #3.
KW Human; kinase; gene therapy; bioreactor; mental disorder;
KW biological disorder.
XX Homo sapiens.
PN W0200123579-A1.
PD 05-APR-2001.
XX 27-SEP-2000; 2000WO-US26621.
XX 28-SEP-1999; 99US-0156511.
XX (LEXI-) LEXICON GENETICS INC.
PI Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
WPI; 2001-266166/27.
DR N-PSDB; AAD03814.

PT New isolated human kinase polynucleotide useful for generating
PT antibodies, as reagents in diagnostic assays and for screening for
PT compounds useful for treating mental, biological or medical diseases -
XX Disclosure; Page 30; 38pp; English.

XX The present sequence is novel human protein
CC (NHP) known as human kinase. The human kinases share structural
CC similarity with animal kinases, more particularly serine or
CC threonine protein kinases. Human kinase cDNA is useful for the
CC detection of mutant human kinase for the diagnosis of disease,
CC and also as a therapeutic. It is useful for screening drugs
CC effective in the treatment of symptomatic or phenotypic
CC manifestations perturbing the normal function of NHP in the
CC body. The NHP nucleotide sequences are useful for generation of
CC antibodies, as reagents in diagnostic assays, for the
CC identification of other cellular gene products related to human
CC kinases, and as reagents in assays for screening compounds that
CC are useful for treating mental, biological or medical disorders.
CC NHP oligonucleotides are used as probes. The labelled NHP probes

CC are useful for screening human genomic library for identifying
CC polymorphisms and as primers in amplification assays to detect
CC mutations within the exons, introns and splice sites that can
CC be used in diagnostics and pharmacogenomics. Nucleotide construct
CC encoding NHP products are used to genetically engineer cells
CC in vivo that functions as bioreactors in the body delivering a
CC continuous supply of NHP to the body. Nucleotide constructs
CC encoding functional NHPs are used in gene therapy for the
CC modulation of NHP expression.

XX
SQ Sequence 324 AA;
Query Match 98.6%; Score 1638; DB 22; Length 324;
Best Local Similarity 100.0%; Pred. No. 7.1e-166;
Matches 310; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEKYEKLAKTGEYSYGVVFKCRNKTSGQVVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MEKYEKLAKTGEYSYGVVFKCRNKTSGQVVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEPNPVGADGVKSVLWQTLQALNFCIH 120
Db 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEPNPVGADGVKSVLWQTLQALNFCIH 120
QY 121 NCIHRIKIPENILITKQGIKICDFGFAQILIPGDAYDYVATRWYRAPELLVGDTCYGS 180
Db 121 NCIHRIKIPENILITKQGIKICDFGFAQILIPGDAYDYVATRWYRAPELLVGDTCYGS 180
QY 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIIPRHQSIFKSNGFHGISIP 240
Db 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIIPRHQSIFKSNGFHGISIP 240
QY 241 EPEDMETLEEKFSVDVHPVALNFMKGLKMPDRLTCSQLLESSYFDSFQEAQIKRKARN 300
Db 241 EPEDMETLEEKFSVDVHPVALNFMKGLKMPDRLTCSQLLESSYFDSFQEAQIKRKARN 300
QY 301 EGRNRRRQ--QVLP 310
Db 301 EGRNRRRQ--QVLP 310

RESULT 4
AAE00494
ID AAE00494 standard; Protein; 347 AA.

XX AAE00494;
XX 19-JUN-2001 (first entry)
XX Human kinase #5.
XX Human; kinase; gene therapy; bioreactor; mental disorder;
KW biological disorder.
XX Homo sapiens.
XX W0200123579-A1.
XX 05-APR-2001.
XX 27-SEP-2000; 2000WO-US26621.
XX 28-SEP-1999; 99US-0156511.
XX (LEXI-) LEXICON GENETICS INC.
PI Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
WPI; 2001-266166/27.
DR N-PSDB; AAD03816.

XX New isolated human kinase polynucleotide useful for generating
PT antibodies, as reagents in diagnostic assays and for screening for

PI compounds useful for treating mental, biological or medical diseases -

XX Claim 2; Page 32-33; 38pp; English.

XX The present sequence is novel human protein.

CC (NHP) known as human kinase. The human kinases share structural

CC similarity with animal kinases. The human kinases share structural

CC threonine protein kinases. Human kinase cDNA is useful for the

CC detection of mutant human kinase for the diagnosis of disease,

CC and also as a therapeutic. It is useful for screening drugs

CC effective in the treatment of symptomatic or phenotypic

CC manifestations perturbing the normal function of NHP in the

CC body. The NHP nucleotide sequences are useful for generation of

CC antibodies, as reagents in diagnostic assays, for the

CC identification of other cellular gene products related to human

CC kinases, and as reagents in assays for screening compounds that

CC are useful for treating mental, biological or medical disorders.

CC NHP oligonucleotides are used as probes. The labelled NHP probes

CC are useful for screening human genomic library for identifying

CC polymorphisms and as primers in amplification assays to detect

CC mutations within the exons, introns and splice sites that can

CC be used in diagnostics and pharmacogenomics. Nucleotide construct

CC encoding NHP products are used to genetically engineer cells

CC in vivo that functions as bioreactors in the body delivering a

CC continuous supply of NHP to the body. Nucleotide constructs

CC encoding functional NHPs are used in gene therapy for the

CC modulation of NHP expression.

XX Sequence 347 AA;

SQ Query Match 98.4%; Score 1636; DB 22; Length 347;
Best Local Similarity 90.8%; Pred. No. 1.3e-165;
Matches 315; Conservative 0; Mismatches 0; Indels 32; Gaps 1;

QY 1 MEKYEKLAKTGEYSYGVVFKCRNKTSGQVAVKKEVESEDDPVVKKIALREIRMLKQLKH 60
|||
Db 1 MEKYEKLAKTGEYSYGVVFKCRNKTSGQVAVKKEVESEDDPVVKKIALREIRMLKQLKH 60

QY 61 PNLVNLIEVFRKRKQHLVFEYCDHTLLNELERNPNGVADGVKSVLWQTLQALNFCIH 120
|||
Db 61 PNLVNLIEVFRKRKQHLVFEYCDHTLLNELERNPNGVADGVKSVLWQTLQALNFCIH 120

QY 121 NCIHREDIKPENILITKQGIKICDFGFAQILIPGDAYTDVATRWYRAPELLVSTQYGS 190
|||
Db 121 NCIHREDIKPENILITKQGIKICDFGFAQILIPGDAYTDVATRWYRAPELLVSTQYGS 180

QY 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTL----- 217
|||
Db 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLVETGFRHVDAQGLELTSDFPA 240

QY 218 -----GKLIPRHQSIFKSNGFHGISIPEPEMETLEEKFSVHVPVAFVNMKGCLK 268
|||
Db 241 VASQSAGITGKLIPIRHQSIFKSNGFHGISIPEPEMETLEEKFSVHVPVAFVNMKGCLK 300

QY 269 MNPDDRLTCSQLLESSYFDSFQEAQIKRKARNEGRNRRRQQLPLKS 315
|||
Db 301 MNPDDRLTCSQLLESSYFDSFQEAQIKRKARNEGRNRRRQQLPLKS 347

RESULT 5
AAU03525
ID AAU03525 standard; Protein; 360 AA.
XX
AC AAU03525;
XX
DT 12-SEP-2001 (first entry)
XX
DE Human protein kinase #25.
XX
KW Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
KW metabolic disorder; immune related disease; neurological disorder;
KW neurodegenerative disorder; inflammatory disorder; infectious disease;
KW reproductive disorder.

XX Homo sapiens.

CS WO200138503-A2.

XX 31-MAY-2001.

XX 22-NOV-2000; 2000WO-US32085.

XX 24-NOV-1999; 99US-0167482.

XX (SUGEN-) SUGEN INC.

XX Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
PI Flanagan P, Clary D;

DR WPI; 2001-343950/36.

DR N-PSDB; AAS06725.

XX Nucleic acids encoding human kinase polypeptides, useful for preventing
PI diagnosing and/or treating e.g. cancer, immune, cardiovascular and
PI neuronal-associated diseases, and microbial infections -

XX Claim 7; Figure 2; 433pp; English.

CC AAU03501-AAU03557 represent novel human protein kinases #1-57. The
CC novel protein kinases have been identified as members of the tyrosine
CC or serine/threonine kinase (PTK and STK) families. The polynucleotides
CC encoding protein kinases and the polypeptides may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate kinase expression. For example, they may be used to treat
CC cancers (especially cancers of haematopoietic origin), cardiovascular
CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
CC immune related diseases (e.g. rheumatoid arthritis), neurological
CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).
CC Additionally, polynucleotides encoding protein kinases may be
CC used for gene therapy and as DNA probes in diagnostic assays.
CC The protein kinase polypeptides may be used as antigens in the production
CC of antibodies against the protein kinases and in assays to identify
CC modulators of protein kinase expression and activity.

XX Sequence 360 AA;

SQ Query Match 97.4%; Score 1619; DB 22; Length 360;
Best Local Similarity 93.4%; Pred. No. 8.9e-164;
Matches 311; Conservative 2; Mismatches 0; Indels 20; Gaps 2;

QY 1 MEKYEKLAKTGEYSYGVVFKCRNKTSGQVAVKKEVESEDDPVVKKIALREIRMLKQLKH 60
|||
Db 1 MEKYEKLAKTGEYSYGVVFKCRNKTSGQVAVKKEVESEDDPVVKKIALREIRMLKQLKH 60

QY 61 PNLVNLIEVFRKRKQHLVFEYCDHTLLNELERNPNGVADGVKSVLWQTLQALNFCIH 120
|||
Db 61 PNLVNLIEVFRKRKQHLVFEYCDHTLLNELERNPNGVADGVKSVLWQTLQALNFCIH 120

QY 121 NCIHREDIKPENILITKQGIKICDFGFAQIL-----IPGDAYTDYVA 162
|||
Db 121 NCIHREDIKPENILITKQGIKICDFGFAQILSNHSHVGRIDLPLDAFVPGDAYTDYVA 180

QY 163 TRWYRAPELLVGTQYSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIP 222
|||
Db 181 TRWYRAPELLVGTQYSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIP 240

QY 223 RHQSIFKSNGFHGISIPEPEMETLEEKFSVHVPVAFVNMKGCLKNPDDRLTCSQLLE 282
|||
Db 241 RHQSIFKSNGFHGISIPEPEMETLEEKFSVHVPVAFVNMKGCLKNPDDRLTCSQLLE 300

QY 283 SSYFDSFQEAQIKRKARNEGRNRRRQ--QVLPL 313
|||
Db 301 SSYFDSFQEAQIKRKARNEGRNRRRQ--QVLPL 333

RESULT 6
AAE00491
ID AAE00491 standard; Protein; 356 AA.
XX
XX AAE00491;
XX
DT 19-JUN-2001 (first entry)
XX
DE Human kinase #2.
XX
KW Human; kinase; gene therapy; bioreactor; mental disorder;
KW biological disorder.
XX
CS Homo sapiens.
XX
PN WO200123579-A1.
XX
PD 05-APR-2001.
XX
PF 27-SEP-2000; 2000WO-US26621.
XX
PR 28-SEP-1999; 99US-0156511.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Donohc G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
XX
DR WPI; 2001-266166/27.
DR N-PSDB; AAD03813.
XX
PF New isolated human kinase polynucleotide useful for generating
PF antibodies, as reagents in diagnostic assays and for screening for
PF compounds useful for treating mental, biological or medical diseases
XX
PS Disclosure; Page 28-29; 38pp; English.
XX
CC The present sequence is novel human protein
CC (NHP) known as human kinase. The human kinases share structural
CC similarity with animal kinases, more particularly serine or
CC threonine protein kinases. Human kinase cDNA is useful for the
CC detection of mutant human kinase for the diagnosis of disease,
CC and also as a therapeutic. It is useful for screening drugs
CC effective in the treatment of symptomatic or phenotypic
CC manifestations perturbing the normal function of NHP in the
CC body. The NHP nucleotide sequences are useful for generation of
CC antibodies, as reagents in diagnostic assays, for the
CC identification of other cellular gene products related to human
CC kinases, and as reagents in assays for screening compounds that
CC are useful for treating mental, biological or medical disorders.
CC NHP oligonucleotides are used as probes. The labelled NHP probes
CC are useful for screening human genomic library for identifying
CC polymorphisms and as primers in amplification assays to detect
CC mutations within the exons, introns and splice sites that can
CC be used in diagnostics and pharmacogenomics. Nucleotide construct
CC encoding NHP products are used to genetically engineer cells
CC in vivo that functions as bioreactors in the body delivering a
CC continuous supply of NHP to the body. Nucleotide constructs
CC encoding functional NHPs are used in gene therapy for the
CC modulation of NHP expression.
XX
SQ Sequence 356 AA;
Query Match 97.0%; Score 1612; DB 22; Length 356;
Best Local Similarity 90.6%; Pred. No. 4.9e-163;
Matches 310; Conservative 0; Mismatches 0; Indels 32; Gaps 1;
QY 1 MEKYEKLAKTGEGSYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMJLQKLKH 60
DB 1 MEKYEKLAKTGEGSYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMJLQKLKH 60
QY 61 PNLVNLIEVFRKRKQHLVFEYCDHTLLNELERNPNGVADGVKSVLWQTLQALNFCCHI 120

DB 61 PNLVNLIEVFRKRKQHLVFEYCDHTLLNELERNPNGVADGVKSVLWQTLQALNFCCHI 120
QY 121 NCIHRIKPENILITKQGIKICDFGFAQILIPGDAYTDYVATRWYRAPELLVGDYQYS 180
DB 121 NCIHRIKPENILITKQGIKICDFGFAQILIPGDAYTDYVATRWYRAPELLVGDYQYS 180
QY 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTL----- 217
DB 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLVETGFRHVDQAGLELLTSSDPPA 240
QY 218 -----GKLIPRHQSIFKSNGFHFGISIPEDMETLEEKFSVDHPVALNFMKGCLK 268
DB 241 VASQSAGITGKLIPRHQSIFKSNGFHFGISIPEDMETLEEKFSVDHPVALNFMKGCLK 300
QY 269 MNPDDRLLTCSQLLESSYFDSFQEAQIKRKARNEGRNRRQV 310
DB 301 MNPDDRLLTCSQLLESSYFDSFQEAQIKRKARNEGRNRRQV 342

RESULT 7
AAB65643
ID AAB65643 standard; Protein; 296 AA.
XX
AC AAB65643;
XX
DT 27-YAR-2001 (first entry)
XX
DE Novel protein kinase, SEQ ID NO: 170.
XX
KW Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
KW immunosuppressive; cardiant; renal; antiinflammatory; antiasthmatic;
KW dermatological; antidiabetic; antifertility; gene therapy; vaccine;
KW immune disorder; cardiovascular disease; neurodegenerative disease;
KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;
KW inflammatory pelvic disease; multiple sclerosis; psoriasis.
XX
CS Mus musculus.
XX
PN WO2000073469-A2.
XX
PD 07-DEC-2000.
XX
PF 26-MAY-2000; 2000WO-US14842.
XX
PR 28-MAY-1999; 99US-0136503.
XX
PA (SUCE-) SUGEN INC.
XX
PI Plowman GD, Martinez R, Whyte D, Sudersanam S;
XX
DR WPI; 2001-032161/04.
DR N-PSDB; AAF44670.
XX
PT Nucleic acids encoding kinase polypeptides, useful for diagnosing and
PT treating immune-related diseases and disorders, cardiovascular disease,
PT neurodegenerative diseases and/or cancers -
XX
PS Claim 10; Fig 1; 310pp; English.
XX
CC The present sequence is a novel protein kinase. The novel protein kinases
CC and the nucleic acids that encode them may be used in the treatment and
CC diagnosis of diseases associated with inappropriate kinase expression
CC such as immune-related diseases and disorders, cardiovascular disease,
CC neurodegenerative diseases and/or cancers. The nucleic acids and
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays. The kinase polypeptides may be used as antigens in the production
CC of antibodies of kinase expression and activity. Anti-kinase antibodies
CC and kinase antagonists may also be used to down regulate kinase
CC expression and activity. Diseases related to kinase expression and
CC activity include rheumatoid arthritis, atherosclerosis, autoimmune
CC disorders, complications of organ transplantation, myocardial infarction,
CC immune disorders, cardiomyopathies, strokes, renal failure,
CC oxidative-stress related disorders, chronic inflammatory bowel disease,

CC Chronic inflammatory pelvic disease, multiple sclerosis, asthma,
CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and
CC reproductive disorders.
XX
SQ Sequence 296 AA;
Query Match 78.0%; Score 1296.5; DB 22; Length 236;
Best Local Similarity 90.0%; Pred. No. 1.7e-129;
Matches 243; Conservative 19; Mismatches 5; Indels 3; Gaps 2;
QY 46 KIALREIRMLKQLKHPNLVNLIEVFRKRKQHLVFEYCDHTLLNELERNPNGVADGVKS 105
DB 1 KIALREIRMLK-LKHPNLVNLIEVFRKRKQHLVFEYCDHTLLNELERNPNGVADGVKS 59
QY 106 VJWQTLQALNFCIHNCIHRDVKPENILITKQGIKICDFGAQILIPGDYTDYVATRW 165
DB 6C VJWQTLQALNFCIHNCIHRDVKPENILITKQGMKICDFGARILIPGDYTDYVATRW 119
QY 166 YRAPELLVGDYQYSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIPRHQ 225
DB 120 YRAPELLVGDYQYSSVDVWAVGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIPRHQ 179
QY 226 SIFKSNGFHGISIPEPEDMETLEKFSVDVHPVALNFMKGLKQNPDDRLLTCSQLLESSY 285
DB 180 SIFRSNQFPGISIPEPEDMETLEKFSNVQPVLSFMKGLKQNPDERLTCAQLLDSAY 239
QY 286 FDSFQEAQIKRKARNEGRNRQ--QVLPL 313
DB 240 FESFQEDQMKRKARSEGRSRRRQQLPL 269
RESULT 8
ABP62954
ID ABP62954 standard; Protein; 358 AA.
XX
AC ABP62954;
XX
DT 14-OCT-2002 (first entry)
XX
DE Human polypeptide SEQ ID NO 391.
XX
KW Human; vulnary; dermatological; neuroprotective; nootropic; cancer;
KW antiparkinsonian; immunostimulant; cytostatic; immunosuppressive;
KW antidiabetic; antiallergic; gene therapy; wound healing; tissue repair;
KW burn; central nervous system disorder; Alzheimer's disease;
KW Parkinson's disease; Huntington's disease; immune disorder;
KW autoimmune disorder; multiple sclerosis; diabetes; allergy.
XX
OS Homo sapiens.
XX
FN WO200218424-A2.
XX
PD 07-MAR-2002.
XX
PF 31-AUG-2001; 2001WO-US27093.
XX
PR 01-SEP-2000; 2000US-0654935.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Asundi V, Zhou P, Xue AJ, Ren F, Zhang J, Wang J;
PI Zhao QA, Wang D, Liu C, Drmanac RT, Wehrman T;
XX
DR WPI; 2002-583321/62.
DR N-PSDB; ABQ93433.
XX
PT New polynucleotide and polypeptides, useful for treatment and diagnosis
PT of Alzheimer's, Parkinson's, Huntington's, amyotrophic lateral
PT sclerosis, immune deficiencies, cancer, autoimmune disorders, multiple
PT sclerosis, diabetes and allergies -
XX
PS Claim 20; SEQ ID NO 391; 284pp + Sequence Listing; English.
XX

CC The invention relates to an isolated polynucleotide (I) comprising one of
CC 245 sequences (ABQ93288-ABQ93532). Treating a condition comprising
CC administering to a mammalian subject a composition comprising the protein
CC (I) encoded by (I) (ABP62809-ABP63053) or an antibody (III) to (II).
CC (I), (II) and (III) are useful for diagnostic evaluation of disorders.
CC (I) is useful for gene therapy of diseases and (II) can be used for
CC therapeutic treatment. Diseases that may be treated include wound healing
CC and tissue repair, burns, central nervous system disorders (e.g.
CC Alzheimer's, Parkinson's, Huntington's and amyotrophic lateral
CC sclerosis), immune deficiencies, cancer, autoimmune disorders, multiple
CC sclerosis, diabetes and allergies.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 358 AA;
Query Match 71.5%; Score 1188.5; DB 23; Length 358;
Best Local Similarity 70.5%; Pred. No. 7.8e-118;
Matches 213; Conservative 42; Mismatches 46; Indels 1; Gaps 2;
QY 1 MEKYEKLAKTGEYSYGVVFKCRNKTSSQVAVVKKFVESEDDPVVKKIALREIRMLKQLKH 60
DB 2 MEKYEKIGKIGESYGVVFKCRNRTGQIVAKKLESEDDPVVKKIALREIRMLKQLKH 61
QY 61 PNLVNLIEVFRKRKQHLVFEYCDHTLLNELERNPNGVADGVKS VJWQTLQALNFCIH 120
DB 62 PNLVNLIEVFRKRRLHLVFEYCDHTVLHELDYQYGVPEHLVKSITWQTLQAVNFC 121
QY 121 NCJHRDIKPENILITKQGIKICDFGAQILI-PGDYTDYVATRWYRAPPELLVGDYQY 179
DB 122 NCJHRDVKPENILITKHSVIKLCDFGARLLTGPSDYTDYVATRWYRSPELLVGDYQY 181
QY 180 SSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIPRHQSIFKSNGFHGISI 239
DB 182 PPVDVWAIGCVFAELLSGVPLWPGKSDVDQLYLIIRTLGDLIPRHQQVFSYFSGVKI 241
QY 240 PEPEDMETLEKFSVDVHPVALNFMKGLKQNPDDRLLTCSQLLESSYFDSFQEAQIKRKAR 299
DB 242 PEPEDMEPLELKFPN-SYPALGLLKGLHMDPTERTLCEQLLHHPYFENIREIEDLAKEH 301
QY 300 NE 301
DB 302 NK 303
RESULT 9
AAB65642
ID AAB65642 standard; Protein; 247 AA.
XX
AC AAB65642;
XX
DT 27-MAR-2001 (first entry)
XX
DE Novel protein kinase, SEQ ID NO: 169.
XX
KW Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
KW immunosuppressive; cardiac; renal; antiinflammatory; antiasthmatic;
KW dermatological; antidiabetic; antiinfertility; gene therapy; vaccine;
KW immune disorder; cardiovascular disease; neurodegenerative disease;
KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;
KW inflammatory pelvic disease; multiple sclerosis; psoriasis.
XX
OS Homo sapiens.
XX
FN WO200073469-A2.
XX
PD 07-DEC-2000.
XX
PF 26-MAY-2000; 2000WO-US14842.
XX
PR 28-MAY-1999; 99US-0136503.
XX

PA (SUGEN-) SUGEN INC.
XX
PI Plowman GD, Martinez R, Whyte D, Sudersanam S;
XX
DR WPI: 2001-032161/04.
DR N-PSDB; AAF44669.
XX
XX Nucleic acids encoding kinase polypeptides, useful for diagnosing and
PT treating immune-related diseases and disorders, cardiovascular disease,
PT neurodegenerative diseases and/or cancers -
XX
PS Claim 10; Fig 1; 310pp; English.
XX
XX The present sequence is a novel protein kinase. The novel protein kinases
CC and the nucleic acids that encode them may be used in the treatment and
CC diagnosis of diseases associated with inappropriate kinase expression;
CC such as immune-related diseases and disorders, cardiovascular disease,
CC neurodegenerative diseases and/or cancers. The nucleic acids and
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays. The kinase polypeptides may be used as antigens in the production
CC of antibodies of kinase expression and activity. Anti-kinase antibodies
CC and kinase antagonists may also be used to down regulate kinase
CC expression and activity. Diseases related to down regulate kinase
CC activity include rheumatoid arthritis, atherosclerosis, autoimmune
CC disorders, complications of organ transplantation, myocardial infarction,
CC immune disorders, cardiomyopathies, strokes, renal failure,
CC oxidative-stress related disorders, chronic inflammatory bowel disease,
CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,
CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and
CC reproductive disorders.
XX
SQ Sequence 247 AA;

Query Match 69.1%; Score 1149; DB 22; Length 247;
Best Local Similarity 98.6%; Pred. No. 7.4e-114;
Matches 217; Conservative 1; Mismatches 0; Indels 2; Gaps 1;

QY 96 NGVADGVTKSVLWQTLQALNFCIHNCIHRDIKPENILITKQGIKICDPFGAQILIPGD 155
DB 1 NGVADGVTKSVLWQTLQALNFCIHNCIHRDIKPENILITKQGIKICDPFGAQILIPGD 60
QY 156 AYTDYVATRWYRAPELLVGDTQYGSSVDIWAIGCVFAELTQPLWPGKSDVDQLYLIR 215
DB 61 AYTDYVATRWYRAPELLVGDTQYGSSVDIWAIGCVFAELTQPLWPGKSDVDQLYLIR 120
QY 216 TLGKLIPRHQSIFKNSGFFHGISIPEPEDMETLEEKFSVDHPVALNFMKGLCNPDRL 275
DB 121 TLGKLIPRHQSIFKNSGFFHGISIPEPEDMETLEEKFSVDHPVALNFMKGLCNPDRL 180
QY 276 TCSQLLESSYFDSFQEAQIKRKARNEGRRRRQ--QVLPL 313
DB 181 TCSQLLESSYFDSFQEAQIKRKARNEGRRRRQQNQLLPL 220

RESULT 10
AAG78547
ID AAG78547 standard; Protein: 228 AA.
XX
AC AAG78547;
XX
DT 08-MAR-2002 (first entry)
XX
DE Human kinase 14257 amino acid sequence.
XX
KW Protein kinase; enzyme; cytostatic; osteopathic; hepatotropic;
KW antidiabetic; neuroprotective; antiarthritic; dermatological;
KW immunosuppressive; antiinflammatory; antithyroid; antipsoriatic;
KW ophthalmological; antiallergic; antiasthmatic; antiatherosclerotic;
KW hypotensive; vasotropic; antiarrhythmic; virucide; anorectic;
KW metabolic; immunomodulator; analgesic; cellular proliferative disorder;
KW cancer; acute lymphoblastic leukaemia; Hodgkin's disease;
KW bone metabolism disorder; osteoporosis; immune system disorder;
KW inflammatory; diabetes mellitus; osteoarthritis; asthma;

KW cardiovascular disorder; hypertension; coronary artery disease;
KW endothelial cell disorder; psoriasis.
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT 4..218
FT /note= "eukaryotic protein kinase domain"
FT Modified-site 9..15
FT /note= "tyrosine kinase phosphorylation site"
FT Modified-site 23..26
FT /note= "N-glycosylation site"
FT Modified-site 27..32
FT /note= "N-myristoylation site"
FT Modified-site 38..41
FT /note= "casein kinase II phosphorylation site"
FT Modified-site 97..102
FT /note= "N-myristoylation site"
FT Modified-site 132..134
FT /note= "serine/threonine protein kinase active-site
FT signal site"
FT 161..218
FT /note= "kinase transferase protein
FT serine/threonine-protein ATP-binding II phosphorylation
FT casein alpha chain domain"
FT Modified-site 180..183
FT /note= "casein kinase II phosphorylation site"
FT Modified-site 188..193
FT /note= "N-myristoylation site"
FT Modified-site 204..211
FT /note= "tyrosine kinase phosphorylation site"
FT Modified-site 205..208
FT /note= "casein kinase II phosphorylation site"
FT
XX WO200179488-A2.
FN 25-OCT-2001.
XX
PF 13-APR-2001; 2001WO-US12188.
XX
PR 13-APR-2000; 2000US-196910P.
XX (MILL-) MILLENNIUM PHARM INC.
PI Kapeller-libermann R;
XX
DR WPI: 2002-034355/04.
DR N-PSDB; AAI64248.
XX
XX New 14257 polypeptides (protein kinases), useful as diagnostic targets
PT and therapeutic agents for controlling cellular proliferative and/or
PT differentiative disorder, bone disorders, immune disorders and
PT cardiovascular disorders -
XX
PS Claim 1c; Fig 1A; 98pp; English.
XX
CC The invention relates to an isolated 14257 polypeptide and nucleic
CC acid encoding it. The 14257 protein is a protein kinase that acts as a
CC modulating agent in regulating a variety of cellular processes,
CC including cell proliferation, differentiation, growth and division.
CC The activity of the protein of the invention may be described as;
CC cytostatic; osteopathic; hepatotropic; antidiabetic; neuroprotective;
CC antiarthritic; dermatological; immunosuppressive; antiinflammatory;
CC antithyroid; antipsoriatic; ophthalmological; antiallergic;
CC antiasthmatic; antiatherosclerotic; hypotensive; vasotropic;
CC antiarrhythmic; virucide; anorectic; metabolic; immunomodulator and
CC analgesic. The protein of the invention may act as a novel diagnostic
CC target or therapeutic agent controlling certain disorders, for example
CC kinase-associated or other 14257-associated disorders. These may include
CC cellular proliferative disorders such as cancers e.g. acute lymphoblastic
CC leukemia or Hodgkin's disease. Other disorders include bone metabolism
CC disorders such as osteoporosis, disorders of the immune system, e.g.
CC inflammatory, diabetes mellitus, osteoarthritis and asthma. Proteins of

CC kinase is a cdc2-related kinase. The KKIAMRE kinase polynucleotides can
CC be used to express recombinant protein for analysis, characterisation or
CC therapeutic use, as markers for tissues in which the protein is
CC preferentially expressed, as molecular weight markers on Southern gels,
CC as chromosome markers or tags, to compare endogenous DNA sequences in
CC patients to identify potential genetic disorders, as probes to hybridise
CC and discover novel related sequences, as a source of PCR primers, and as
CC an antigen to induce anti-DNA antibodies. The polypeptides can be used
CC in assay to discover biological activity, to raise antibodies, as tissue
CC markers, and to isolate correlative receptors or ligands. The
CC polynucleotides may also be used for gene therapy for the treatment of
CC disorders which are mediated by KKIAMRE kinase.

XX Sequence 566 AA;
SQ Query Match 58.2%; Score 967.5; DB 21; Length 566;
Best Local Similarity 54.6%; Pred. No. 6e-94;
Matches 177; Conservative 67; Mismatches 65; Indels 15; Gaps 5;

QY 1 MEKYEKLAKTGEGSYGVVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIRMLKQLKH 60
DB 1 MEKYENLGLVGEGSYGMVKCRNKDSGRIVAIAKKFESDDDDKAVKKIAMREIKLKQLRH 60
QY 61 PNLVNLIEVFRRRKRMHLVFEYCDHTLLNELEPNPVGADGVIKSVLWQTLQALNECHIH 120
DB 61 ENLVNLLEVCKKKRWLVFEPVDHTILDOLFPNGLDDQVVKYLFQINGIGFCHSH 120
QY 121 NCIHRIKPENILITKQGIKIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDTOYG 179
DB 121 NIHRDIKPENILVSQSGVVKLCDFGFARTLAAGEVYTDYVATRWYRAPELLVGDVXYG 180
QY 180 SSVDIWAIGCVFAELLTGQPLWFGKSDVDQLYLIIRTLGKLIPRHQSIFKSNGFPHGISI 239
DB 181 KAVDVWAIGCLVTEMLMGBPLFPGCSDDIDQLYLIMRCLGNLIPRHQELFYKNPVFAGVRL 240
QY 240 PEPEDMETLEKFSVDHVFVALNFKGCLKNPDDRLTCSLLESSYF--DSF-----QEA 292
DB 241 PEIKESEPLERRYPKLSEVVIDLAKKCLHVDPDKRPPFCAELLHHDFQMDGFAERFSQEL 300
QY 293 QIK--RKARN-----EGNRNRRCQ 309
DB 301 QMKVQKDARNISLSKKSQNRKKEK 324

RESULT 13
ABU11689
ID ABU11689 standard; Protein; 197 AA.
XX
AC ABU11689;
XX

DT 13-FEB-2003 (first entry)
DE Human MDDT polypeptide SEQ ID 636.
XX

KW MDDT; human; disease detection and treatment molecule polypeptide;
KW anti-inflammatory; immunosuppressive; osteopathic; cytostatic; anti-HIV;
KW haemostatic; nephrotropic; antianaemic; antipsoriatic; hepatotropic;
KW gene therapy; protein replacement therapy; cell proliferative disorder;
KW cancer; adenocarcinoma; leukaemia; lymphoma; melanoma; myeloma; sarcoma;
KW anaemia; Crohn's disease; acquired immunodeficiency syndrome; AIDS;
KW Goodpasture's syndrome; inflammation; osteoporosis; thrombocytopaenia;
KW psoriasis; hepatitis.

XX Homo sapiens.
OS
XX
PN WO200279449-A2.
XX

PD 10-OCT-2002.

XX 27-MAR-2002; 2002WO-US09944.

PF 28-MAR-2001; 2001US-279619P.

PR 29-MAR-2001; 2001US-280067P.

PR 23-MAR-2001; 2001US-280068P.
PR 16-MAY-2001; 2001US-291280P.
PR 17-MAY-2001; 2001US-291829P.
PR 17-MAY-2001; 2001US-291849P.
PR 19-JUN-2001; 2001US-299428P.
PR 20-JUN-2001; 2001US-299776P.
PR 20-JUN-2001; 2001US-300001P.
XX (INCY-) INCYTE GENOMICS INC.
PA
XX Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
PI Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshey SR;
PI Daugherty SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstin EH;
PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;
XX
DR WPI; 2003-058431/05.
N-PSDB; ABX34679.

XX New purified disease detection and treatment molecule proteins and
PT polynucleotides, useful for diagnosing, treating or preventing cancers
PT (e.g. leukemia or sarcoma), anemia, Crohn's disease, AIDS, osteoporosis
PT or hepatitis -
XX
PS Claim 27; SEQ ID NO 636; 339pp + Sequence Listing; English.
XX

CC This invention describes a novel disease detection and treatment molecule
CC polypeptide (MDDT) which has anti-inflammatory, immunosuppressive,
CC osteopathic, cytostatic, anti-HIV, haemostatic, nephrotropic,
CC antianaemic, antipsoriatic and hepatotropic activity. The polynucleotides
CC and the polypeptides of the invention can be used for gene therapy,
CC protein replacement therapy and are useful for treating a variety of
CC diseases or conditions. These polypeptides or polynucleotides are
CC particularly useful for diagnosing, treating or preventing cell
CC proliferative disorders (e.g. cancers including adenocarcinoma,
CC leukaemia, lymphoma, melanoma, myeloma or sarcoma), anaemia, Crohn's
CC disease, acquired immunodeficiency syndrome (AIDS), Goodpasture's
CC syndromes, inflammation, osteoporosis, thrombocytopaenia, psoriasis or
CC hepatitis. ABU11450-ABU11845 represent the MDDT polynucleotides encoded
CC by ABU11450-ABU11845, described in the disclosure of the invention.

CC NOTE: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 197 AA;

Query Match 57.0%; Score 948; DB 24; Length 197;
Best Local Similarity 98.9%; Pred. No. 1.5e-92;
Matches 180; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MEKYEKLAKTGEGSYGVVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIRMLKQLKH 60
DB 6 MEKYEKLAKTGEGSYGVVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIRMLKQLKH 65

QY 61 PNLVNLIEVFRRRKRMHLVFEYCDHTLLNELEPNPVGADGVIKSVLWQTLQALNPFCHIH 120
DB 66 PNLVNLIEVFRRRKRMHLVFEYCDHTLLNGLERNPVGADGVIKSVLWQTLQALNPFCHIH 125

QY 121 NCIHRIKPENILITKQGIKIKICDFGFAQILIPGDAYTDYVATRWYRAPELLVGDTOYGS 180
DB 126 NCIHRIKPENILITKQGIKIKICDFGFAQILIPGDAYTDYATRWYRAPELLVGDTOYGS 185

QY 181 SV 182
DB 186 SV 187

RESULT 14
ABP96087
ID ABP96087 standard; Protein; 362 AA.
XX
AC ABP96087;
XX

Qy 1 MEKYEKLAKTGEGSYGVVFKCRNKTSGQWAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MEKYEKLAKTGEGSYGVVFKCRNKTSGQWAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Qy 61 PNLVNLIEVFRRRKRMHLVFEYCDHTLLNELEPNPVGADGVKSVLWQTLQALNFCIH 120
Db 61 PNLVNLIEVFRRRKRMHLVFEYCDHTLLNELEPNPVGADGVKSVLWQTLQALNFCIH 120
Qy 121 NCIHRRDIKPENILITKQGIKICDFGFAQIL 151
Db 121 NCIHRRDIKPENILITKQGIKICDFGFAQIL 151

Search completed: November 13, 2003, 14:18:48
Job time : 37.1631 secs

Db 1 MENFQVKEIGEGTYGVVYKARNKLTGEVVALKKIRLDTETEGVPSTAIRESLLKELNH 60
QY 61 PNLVNLIEVFRKRKMHLVFEYCDHTLLNELRNP-NGVADGVKSVLWQTLQALNFCHI 119
Db 61 PNLVNLIEVFRKRKMHLVFEYCDHTLLNELRNP-NGVADGVKSVLWQTLQALNFCHI 119
Db 61 PNLVNLIEVFRKRKMHLVFEYCDHTLLNELRNP-NGVADGVKSVLWQTLQALNFCHI 119
QY 120 HNCIHRDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDTOY 178
Db 121 HRVHLRDLKPNQLLINTEGAIKLADFGARAFGVPRVTYTHEVVTLWYRAPEILLGSKYY 180
QY 179 GSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIIPRHQSIFKSNNGFFHGI- 237
Db 181 STAVDIWSLGCIFAEMVTRRALFPGDSEIDQLFRIFRTLGT-----PDEVVWPGVT 231
QY 238 SIPE--PEDMETLEKEFSDVHPV----ALNFMKGCLKMNPPDDLTCSQLLESSYF 286
Db 232 SMPDYKPSFPKWARQDFSKVVPPLDEDRSLLSQMLHYDPNKRISAKAALAHPPF 286

RESULT 5
US-08-969-106-2
; Sequence 2, Application US/08969106
; Patent No. 5986055
; GENERAL INFORMATION:
; APPLICANT: Yang, M.
; APPLICANT: Nandabalan, K.
; APPLICANT: Schulz, V.
; TITLE OF INVENTION: CDK2 INTERACTIONS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/969,106
; FILING DATE: 13-NOV-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7934-057
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 298 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-969-106-2

Query Match 34.3%; Score 570; DB 2; Length 298;
Best Local Similarity 40.0%; Pred. No. 8.5e-52;
Matches 118; Conservative 60; Mismatches 99; Indels 18; Gaps 6;
QY 1 MEKYEKLAKTGECSYGWVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MENFQVKEIGEGTYGVVYKARNKLTGEVVALKKIRLDTETEGVPSTAIRESLLKELNH 60
QY 61 PNLVNLIEVFRKRKMHLVFEYCDHTLLNELRNP-NGVADGVKSVLWQTLQALNFCHI 119

Db 61 PNLVNLIEVFRKRKMHLVFEYCDHTLLNELRNP-NGVADGVKSVLWQTLQALNFCHI 120
QY 120 HNCIHRDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDTOY 178
Db 121 HRVHLRDLKPNQLLINTEGAIKLADFGARAFGVPRVTYTHEVVTLWYRAPEILLGSKYY 180
QY 179 GSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIIPRHQSIFKSNNGFFHGI- 237
Db 181 STAVDIWSLGCIFAEMVTRRALFPGDSEIDQLFRIFRTLGT-----PDEVVWPGVT 231
QY 238 SIPE--PEDMETLEKEFSDVHPV----ALNFMKGCLKMNPPDDLTCSQLLESSYF 286
Db 232 SMPDYKPSFPKWARQDFSKVVPPLDEDRSLLSQMLHYDPNKRISAKAALAHPPF 286
RESULT 6
US-09-093-522-25
; Sequence 25, Application US/09093522
; Patent No. 6015700
; GENERAL INFORMATION:
; APPLICANT: Limper, Andrew H.
; APPLICANT: Leof, Edward B.
; APPLICANT: Thomas, Charles F.
; APPLICANT: Gustafson, Michael P.
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
; TITLE OF INVENTION: CARINII
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C., P.A.
; STREET: 60 South Sixth Street, Suite 3300
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/093,522
; FILING DATE: 08-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/874,347
; FILING DATE: 13-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ellinger, Mark S.
; REGISTRATION NUMBER: 34,812
; REFERENCE/DOCKET NUMBER: 07039/055002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-335-5070
; TELEFAX: 612-288-9696
; TELEX:
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 298 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-093-522-25

Query Match 34.3%; Score 570; DB 3; Length 298;
Best Local Similarity 40.0%; Pred. No. 8.5e-52;
Matches 118; Conservative 60; Mismatches 99; Indels 18; Gaps 6;
QY 1 MEKYEKLAKTGECSYGWVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MENFQVKEIGEGTYGVVYKARNKLTGEVVALKKIRLDTETEGVPSTAIRESLLKELNH 60
QY 61 PNLVNLIEVFRKRKMHLVFEYCDHTLLNELRNP-NGVADGVKSVLWQTLQALNFCHI 119

Db 61 PNIKLLDVIHTENKLYLVPEFLHQDLKKFMDASALTGIPLPKLSYLFQILQGLAFCHS 120
QY 120 HNCIHRDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDTCY 178
Db 121 HRVLRDLKPQNLLINTEGAIKLADFGLARAFGVPVRTYTHEVVTLMYRAPEILLGSKYY 180
QY 179 GSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIIPRHQSIFKSNNGFFHGI- 237
Db 181 STAVDIWSLGCIFAEMVTRRALFPGDSEIDQLFRIFRTLGT-----PDEVWPGVT 231
QY 238 SIPE--PEDMETLEEKFSVDVHPV---ALNFMKGCLKMNPDRLTCSQLLESSYF 286
Db 232 SMPDYKPSFPKWARKQDFSKVVPPLDEDEGRSJLSQMLHYDPNKRISAKAALAHPPF 286

RESULT 7
US-09-338-125-2
; Sequence 2, Application US/09338125
; Patent No. 6521412
; GENERAL INFORMATION:
; APPLICANT: Yang, M.
; APPLICANT: Nandabalan, K.
; APPLICANT: Schulz, V.
; TITLE OF INVENTION: CDK2 INTERACTIONS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/338,125
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA: US/08/969,106
; APPLICATION NUMBER: 13-NOV-1997
; FILING DATE: 13-NOV-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Wisrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7934-057
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 298 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-09-338-125-2

Query Match 34.3%; Score 570; DB 4; Length 298;
Best Local Similarity 40.0%; Pred. No. 8.5e-52;
Matches 118; Conservative 60; Mismatches 99; Indels 18; Gaps 6;
QY 1 MEKYEKLAKTGEISYGVVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MENFQKVEKIGEGTYGVVYKARNKLTGEVVALKKIRLDTETEGVSPSTAIRISLLKELNH 60
QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEARNP-NGVADGVKSVLWOTLQALNFCHI 119
Db 61 PNIKLLDVIHTENKLYLVFEFLHQDLKKFMDASALTGIPLPKLSYLFQILQGLAFCHS 120

QY 120 HNCIHRDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDTCY 178
Db 121 HRVLRDLKPQNLLINTEGAIKLADFGLARAFGVPVRTYTHEVVTLMYRAPEILLGSKYY 180
QY 179 GSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIIPRHQSIFKSNNGFFHGI- 237
Db 181 STAVDIWSLGCIFAEMVTRRALFPGDSEIDQLFRIFRTLGT-----PDEVWPGVT 231
QY 238 SIPE--PEDMETLEEKFSVDVHPV---ALNFMKGCLKMNPDRLTCSQLLESSYF 286
Db 232 SMPDYKPSFPKWARKQDFSKVVPPLDEDEGRSJLSQMLHYDPNKRISAKAALAHPPF 286
RESULT 8
US-09-266-225D-14
; Sequence 14, Application US/09266225D
; Patent No. 6573364
; GENERAL INFORMATION:
; APPLICANT: Nandabalan, Krishan
; APPLICANT: Kingsmore, Stephen
; APPLICANT: Tchernev, Velizar
; TITLE OF INVENTION: Isolation and Characterization of Hermansky-Pudlak
; TITLE OF INVENTION: Syndrome (HPS) Protein Complexes and Hps Protein-
; TITLE OF INVENTION: Interacting Proteins
; FILE REFERENCE: 15966-523
; CURRENT APPLICATION NUMBER: US/09/266,225D
; CURRENT FILING DATE: 1999-03-10
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-266-225D-14

Query Match 34.3%; Score 570; DB 4; Length 298;
Best Local Similarity 40.0%; Pred. No. 8.5e-52;
Matches 118; Conservative 60; Mismatches 99; Indels 18; Gaps 6;
QY 1 MEKYEKLAKTGEISYGVVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MENFQKVEKIGEGTYGVVYKARNKLTGEVVALKKIRLDTETEGVSPSTAIRISLLKELNH 60
QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEARNP-NGVADGVKSVLWOTLQALNFCHI 119
Db 61 PNIKLLDVIHTENKLYLVFEFLHQDLKKFMDASALTGIPLPKLSYLFQILQGLAFCHS 120
QY 120 HNCIHRDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDTCY 178
Db 121 HRVLRDLKPQNLLINTEGAIKLADFGLARAFGVPVRTYTHEVVTLMYRAPEILLGSKYY 180
QY 179 GSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIIPRHQSIFKSNNGFFHGI- 237
Db 181 STAVDIWSLGCIFAEMVTRRALFPGDSEIDQLFRIFRTLGT-----PDEVWPGVT 231
QY 238 SIPE--PEDMETLEEKFSVDVHPV---ALNFMKGCLKMNPDRLTCSQLLESSYF 286
Db 232 SMPDYKPSFPKWARKQDFSKVVPPLDEDEGRSJLSQMLHYDPNKRISAKAALAHPPF 286

RESULT 9
US-09-417-197-113
; Sequence 113, Application US/09417197
; Patent No. 6518021
; GENERAL INFORMATION:
; APPLICANT: Ole THASTRUP, et al.
; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An
; TITLE OF INVENTION: On A Cellular Response
; FILE REFERENCE: 3759-0110P
; CURRENT APPLICATION NUMBER: US/09/417,197
; CURRENT FILING DATE: 1999-10-07
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.0

```
; SEQ ID NO 113
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: CDK2-EGFP fusion
US-09-417-197-113

Query Match      34.3%; Score 570; DB 4; Length 544;
Best Local Similarity 40.0%; Pred. No. 2.1e-51;
Matches 118; Conservative 60; Mismatches 99; Indels 18; Gaps 6;

QY 1 MEKYEKLAKTGE GSYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
D5 1 MENFQKEKIGEGTYGVVYKARNKLTGEVVALKKIRLDTETEGVPSTAIRESLSLKLNLH 60

QY 61 PNLVNLIEVFRKRKMHLVFEYCDHTLLNELEARNP-NGVADGVKSVLWQTLQALNFCHI 119
D5 61 PNLVNLIEVFRKRKMHLVFEYCDHTLLNELEARNP-NGVADGVKSVLWQTLQALNFCHI 119

D5 61 PNLVNLIEVFRKRKMHLVFEYCDHTLLNELEARNP-NGVADGVKSVLWQTLQALNFCHI 119
D5 61 PNLVNLIEVFRKRKMHLVFEYCDHTLLNELEARNP-NGVADGVKSVLWQTLQALNFCHI 119

QY 120 HNCIHRDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDTOY 178
D5 120 HNCIHRDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDTOY 178

D5 121 HRVLRDLKPQNLLINTEGAIKLADFGLARAFGVPRVTYTHEVVTWYRAPELLVGSKY 183
D5 121 HRVLRDLKPQNLLINTEGAIKLADFGLARAFGVPRVTYTHEVVTWYRAPELLVGSKY 183

QY 179 GSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIIPRHQSIFKSNGFHGI- 237
D5 179 GSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIIPRHQSIFKSNGFHGI- 237

D5 181 STAVDIWSLGCIFAEMVTRRALFPGDSEIDQLFRIFRTLTGT-----PDEVVWPGVT 231
D5 181 STAVDIWSLGCIFAEMVTRRALFPGDSEIDQLFRIFRTLTGT-----PDEVVWPGVT 231

QY 238 SIPE--PEDMETLEEKFSVDHPV----ALNFMKGCLKMNPDRLTCSQLLESSYF 286
D5 238 SIPE--PEDMETLEEKFSVDHPV----ALNFMKGCLKMNPDRLTCSQLLESSYF 286

D5 232 SMPDYKPSFPKWARQDFSKVVPPLDEDEGRSLLSQMLHYDPNKRISAKAALAHPPF 286
D5 232 SMPDYKPSFPKWARQDFSKVVPPLDEDEGRSLLSQMLHYDPNKRISAKAALAHPPF 286

RESULT 10
US-09-417-197-115
; Sequence 115, Application US/09417197
; Patent No. 6518021
; GENERAL INFORMATION:
; APPLICANT: Ole THASTRUP, et al.
; TITLE OF INVENTION: A Method For Extracting Quantitative information Relating To An
; FILE REFERENCE: 3759-0110P
; CURRENT APPLICATION NUMBER: US/09/417,197
; CURRENT FILING DATE: 1999-10-07
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 115
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: EGFP-CDK2 fusion
US-09-417-197-115

Query Match      34.3%; Score 570; DB 4; Length 544;
Best Local Similarity 40.0%; Pred. No. 2.1e-51;
Matches 118; Conservative 60; Mismatches 99; Indels 18; Gaps 6;

QY 1 MEKYEKLAKTGE GSYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
D5 247 MENFQKEKIGEGTYGVVYKARNKLTGEVVALKKIRLDTETEGVPSTAIRESLSLKLNLH 306

QY 61 PNLVNLIEVFRKRKMHLVFEYCDHTLLNELEARNP-NGVADGVKSVLWQTLQALNFCHI 119
D5 61 PNLVNLIEVFRKRKMHLVFEYCDHTLLNELEARNP-NGVADGVKSVLWQTLQALNFCHI 119

D5 307 PNLVNLIEVFRKRKMHLVFEYCDHTLLNELEARNP-NGVADGVKSVLWQTLQALNFCHI 366
D5 307 PNLVNLIEVFRKRKMHLVFEYCDHTLLNELEARNP-NGVADGVKSVLWQTLQALNFCHI 366

QY 120 HNCIHRDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDTOY 178
D5 120 HNCIHRDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDTOY 178

D5 367 HRVLRDLKPQNLLINTEGAIKLADFGLARAFGVPRVTYTHEVVTWYRAPELLVGSKY 426
D5 367 HRVLRDLKPQNLLINTEGAIKLADFGLARAFGVPRVTYTHEVVTWYRAPELLVGSKY 426

QY 179 GSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIIPRHQSIFKSNGFHGI- 237
D5 179 GSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIIPRHQSIFKSNGFHGI- 237

D5 427 STAVDIWSLGCIFAEMVTRRALFPGDSEIDQLFRIFRTLTGT-----PDEVVWPGVT 477
D5 427 STAVDIWSLGCIFAEMVTRRALFPGDSEIDQLFRIFRTLTGT-----PDEVVWPGVT 477
```

```
QY 238 SIPE--PEDMETLEEKFSVDHPV----ALNFMKGCLKMNPDRLTCSQLLESSYF 286
D5 478 SMPDYKPSFPKWARQDFSKVVPPLDEDEGRSLLSQMLHYDPNKRISAKAALAHPPF 532

RESULT 11
US-09-411-628-13
; Sequence 13, Application US/09411628
; Patent No. 6428994
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; TITLE OF INVENTION: CDNA, GENOMIC, AND PREDICTED PROTEIN
; FILE REFERENCE: 13761-707
; CURRENT APPLICATION NUMBER: US/09/411,628
; CURRENT FILING DATE: 1999-10-01
; EARLIER APPLICATION NUMBER: US 60/102,906
; EARLIER FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(298)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-411-628-13

Query Match      34.2%; Score 569; DB 4; Length 298;
Best Local Similarity 40.0%; Pred. No. 1.1e-51;
Matches 118; Conservative 60; Mismatches 99; Indels 18; Gaps 6;

QY 1 MEKYEKLAKTGE GSYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
D5 1 MENFQKEKIGEGTYGVVYKARNKLTGEVVALKKIRXDTETEGVPSTAIRESLSLKLNLH 60

QY 61 PNLVNLIEVFRKRKMHLVFEYCDHTLLNELEARNP-NGVADGVKSVLWQTLQALNFCHI 119
D5 61 PNLVNLIEVFRKRKMHLVFEYCDHTLLNELEARNP-NGVADGVKSVLWQTLQALNFCHI 119

D5 61 PNLVNLIEVFRKRKMHLVFEYCDHTLLNELEARNP-NGVADGVKSVLWQTLQALNFCHI 119
D5 61 PNLVNLIEVFRKRKMHLVFEYCDHTLLNELEARNP-NGVADGVKSVLWQTLQALNFCHI 119

QY 120 HNCIHRDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDTOY 178
D5 120 HNCIHRDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDTOY 178

D5 121 HRVLRDLKPQNLLINTEGAIKLADFGLARAFGVPRVTYTHEVVTWYRAPELLVGSKY 183
D5 121 HRVLRDLKPQNLLINTEGAIKLADFGLARAFGVPRVTYTHEVVTWYRAPELLVGSKY 183

QY 179 GSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIIPRHQSIFKSNGFHGI- 237
D5 179 GSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIIPRHQSIFKSNGFHGI- 237

D5 181 STAVDIWSLGCIFAEMVTRRALFPGDSEIDQLFRIFRTLTGT-----PDEVVWPGVT 231
D5 181 STAVDIWSLGCIFAEMVTRRALFPGDSEIDQLFRIFRTLTGT-----PDEVVWPGVT 231

QY 238 SIPE--PEDMETLEEKFSVDHPV----ALNFMKGCLKMNPDRLTCSQLLESSYF 286
D5 232 SMPDYKPSFPKWARQDFSKVVPPLDEDEGRSLLSQMLHYDPNKRISAKAALAHPPF 286

RESULT 12
US-08-874-347-26
; Sequence 26, Application US/08874347
; Patent No. 5863741
; GENERAL INFORMATION:
; APPLICANT: Limper, Andrew H.
; APPLICANT: Leof, Edward B.
; APPLICANT: Thomas, Charles F.
; APPLICANT: Gustafson, Michael P.
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C., P.A.
; STREET: 60 South Sixth Street, Suite 3300
; CITY: Minneapolis
; STATE: MN
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 14:16:37 ; Search time 13.3233 Seconds
(without alignments)
2273.702 Million cell updates/sec

Title: US-09-671-050-12
Perfect score: 1662
Sequence: 1 MEKYEKLAKTGECSYGVVFK.....RKARNEGRNRRRQQLPLKS 315

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*

1:	pir1:*
2:	pir2:*
3:	pir3:*
4:	pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1183.5	71.2	358	2 S23383	protein kinase (EC
2	1175.5	70.7	376	2 S22745	serine/threonine p
3	581	35.0	305	2 S23382	protein kinase (EC
4	574.5	34.6	292	2 S40021	protein kinase (EC
5	571.5	34.4	294	1 A40444	protein kinase (EC
6	571.5	34.4	294	2 B40444	protein kinase (EC
7	570.5	34.3	302	1 I50474	protein kinase (EC
8	568.5	34.2	294	2 S22440	protein kinase (EC
9	567	34.1	298	2 A41227	protein kinase (EC
10	565.5	34.0	297	2 A37871	protein kinase (EC
11	561.5	33.8	294	2 S23095	protein kinase (EC
12	559	33.6	302	1 A44349	protein kinase (EC
13	557	33.5	297	1 S24913	protein kinase (EC
14	556.5	33.5	294	2 T49271	protein kinase (EC
15	556	33.5	297	2 A29539	protein kinase (EC
16	556	33.5	297	2 I45977	cyclin-dependent k
17	556	33.5	302	2 B44349	protein kinase (EC
18	554	33.3	301	1 S19209	protein kinase (EC
19	553	33.3	298	1 A44878	protein kinase (EC
20	551.5	33.2	311	2 S36619	protein kinase (EC
21	551	33.2	288	1 S42566	protein kinase (EC
22	550.5	33.1	294	1 JQ2243	protein kinase (EC
23	550.5	33.1	294	1 S57928	protein kinase (EC
24	550.5	33.1	294	1 S42049	protein kinase (EC
25	547.5	32.9	303	1 S06011	protein kinase (EC
26	547.5	32.9	308	1 S53538	protein kinase (EC
27	546.5	32.9	297	1 S12009	protein kinase cdc
28	545	32.8	292	1 A46365	protein kinase (EC
29	544.5	32.8	291	2 S23386	protein kinase (EC

30	544	32.7	293	2 JE0374	cyclin-dependent k
31	543	32.7	292	1 I49592	protein kinase (EC
32	543	32.7	292	1 A45091	protein kinase (EC
33	543	32.7	297	1 A36074	protein kinase (EC
34	543	32.7	301	1 S42101	protein kinase (EC
35	541.5	32.6	294	1 S31332	protein kinase (EC
36	540	32.5	296	2 S24386	protein kinase (EC
37	539	32.4	302	1 OKBY85	protein kinase PHO
38	535.5	32.2	302	2 T17115	protein kinase cdc
39	530.5	31.9	346	2 A54820	CDK-activating pro
40	530	31.9	294	2 S51008	protein kinase (EC
41	528	31.8	292	2 S22441	protein kinase (EC
42	527.5	31.7	346	2 A56231	MO15/CDK-activatin
43	527	31.7	346	1 I78840	protein kinase (EC
44	525	31.6	332	1 S41003	protein kinase (EC
45	525	31.6	346	1 I48157	protein kinase (EC

ALIGNMENTS

RESULT 1

S23383
protein kinase (EC 2.7.1.37; cdc2-related KKIALRE - human
C;Species: Homo sapiens (man)
C;Date: 15-Oct-1994 #sequence_revision 15-Oct-1994 #text_change 18-Jun-1999
C;Accession: S23383; S22744
R;Meyerson, M.; Enders, G.H.; Wu, C.L.; Su, L.K.; Gorka, C.; Nelson, C.; Harlow, E.; T
EMBO J. 11, 2909-2917, 1992
A;Title: A family of human cdc2-related protein kinases.
A;Reference number: S23392; MUID:92347325; PMID:1639063
A;Accession: S23383
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-358 <MEY>
A;Cross-references: EMBL:X66358; NID:g36614; PIDN:CAA47002.1; PID:g36615
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
F;3-278/Domain: protein kinase homology <Kin>
F;11-19/Region: protein kinase ATP-binding motif
F;34,52,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match		71.2%;	Score 1183.5;	DB 2;	Length 358;
Best Local Similarity		71.8%;	Pred. No. 3.9e-49;		
Matches 211;		Conservative 41;	Mismatches 41;	Indels 1;	Gaps 1;
QY	1	MEKYEKLAKTGECSYGVVFKCRNKTSGQWAVKKFVESEDDPVVKKIALREIRMLKQLKH	60		
DB	2	MEKYEKIGKIGEGSYGVVFKCRNRTGQIVAIFKFESEDDPVVKKIALREIRMLKQLKH	61		
QY	61	PNLVNLIIEVFRKRKRLHLVFEYCDHTLLNELERNPNGVADGVKISVLWQTLQALNFCHIH	120		
DB	62	PNLVNLIIEVFRKRRLHLVFEYCDHTVLHELDYRQGVPEHLVKSITWQTLQAVNFCHKH	121		
QY	121	NCIHRDIKPNILITKQGIKICDFGFAQILI-PGDAYTDYVATRWYRAPPELLVGDQYQ	179		
DB	122	NCIHRDVKPNILITKHSVIKLCDFGFAQLLTGPSDYTYDVATRWYRSPELLVGDQYQ	181		
QY	180	SSVDIWAIGCVFAELLTGQPLWPKSDVDQLYLIIRTLGKLIIPRHQSIKNSGFFHGISI	239		
DB	182	PPVDVWAIGCVFAELLSGVPLWPKSDVDQLYLIIRKTLGDLIPRHQQVFTNQYFSGVKI	241		
QY	240	PEPEDMETLEKFSVDHPVALNPMKGCGLKMNPDRLTCSQLLESSYFDSFQEAQ	293		
DB	242	PDPEMPELKLKFPNISYPALGLLKGCLHMDPTERTLTGCLLHHPYFENIREIE	295		

RESULT 2

S22745
serine/threonine protein kinase KKIALRE (EC 2.7.1.-) - human
C;Species: Homo sapiens (man)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 19-Dec-1997
C;Accession: S22745

R:Meyerson, M.D.
Submitted to the EMBL Data Library, May 1992
A:Reference number: S22743
A:Accession: S22745
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-376 <MEY>
A:Cross-references: EMBL:X66359
C:Genetics:
A:Introns: 152/3; 170/3
C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
F:3-296/Domain: protein kinase homology <KIN>
F:11-19/Region: protein kinase ATP-binding motif

Query Match 70.7%; Score 1175.5; DB 2; Length 376;
Best Local Similarity 67.6%; Pred. No. 9.5e-49;
Matches 211; Conservative 41; Mismatches 41; Indels 19; Gaps 1;

Qy 1 MEKYEKLAKTGE GSYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQKH 60
Db 2 MEKYEKIGKIGEGSYGVVFKCRNRTGQIVAIKKFLESEDDPVIKKIALREIRMLKQKH 61

Qy 61 PNLVNLIEVFRKRKMHLVFEYCDHTLLNELERNPNGVADGVKSVLWQTLQALNFC 120
Db 62 PNLVNLIEVFRKRRLHLVFEYCDHTVLHEDRYQRGVPEHLVKSITWQTLQAVNFC 121

Qy 121 NCIHRTKPNILITKQGIKICDFGFAQIL-----PGDAYTDYV 161
Db 122 NCIHRTKPNILITKHSVTKLDFGFAQLLXFQPSAAVCFPCSIIGTGPDSYTDYV 161

Qy 162 ATRWYRAPELLVGDTOYGSSVDIWAIGCVFAELLTGQPLWPGKSDVDQYLIIRTLGKLI 221
Db 182 ATRWYRSPPELLVGDTOYGPFPVWVAIGCVFAELLSGVPLWPGKSDVDQYLIIRTLGDLI 241

Qy 222 PRHQSIKSNNGFFHGISPEPEMETLEEKFSVDVHPVALNFMKGCLKMNPDRLTCSQLL 281
Db 242 PRHQVFSTNQYFSGVKIPDPEDMEPLELKFNPISYPALGGLKGCLHMDPTERTCEQLL 301

Qy 282 ESSYFDSFQEAQ 293
Db 302 HHPYFENIREIE 313

RESULT 3
S23382
Protein kinase (EC 2.7.1.37) cdk3 - human
C:Species: Homo sapiens (man)
C:Date: 18-Jun-1993 #sequence_revision 18-Jun-1993 #text_change 18-Jun-1999
C:Accession: S23382; S22743
R:Meyerson, M.; Enders, G.H.; Wu, C.L.; Su, L.K.; Gorla, C.; Nelson, C.; Harlow, E.; Tsai
EMBO J. 11, 2909-2917, 1992
A:Title: A family of human cdc2-related protein kinases.
A:Reference number: S23382; MUID:92347325; PMID:1639363
A:Accession: S23382
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-305 <MEY>
A:Cross-references: EMBL:X66357; NID:G36612; PIDN:CAA47001.1; PID:G36613
C:Genetics:
A:Gene: GDB:CDK3
A:Cross-references: GDB:283456
A:Map position: 12q13-12q13
C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein kinase
F:2-255/Domain: protein kinase homology <KIN>
F:10-18/Region: protein kinase ATP-binding motif
F:33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 35.0%; Score 581; DB 2; Length 305;
Best Local Similarity 39.8%; Pred. No. 5.7e-21;
Matches 123; Conservative 66; Mismatches 96; Indels 24; Gaps 8;

Qy 1 MEKYEKLAKTGE GSYGVVFKCRNKTSGQVAVKKF---VESEDDPVVKKIALREIRMLKQ 57
Db 1 MDMFQKVEKIGEGTYGVVYKAKNRETQGLVALKKIRLDLEMEGVP---STAIREISLLKE 57

Qy 58 LKHPNLVNLIEVFRKRKMHLVFEYCDHTLLNELERNPNGVADGVKSVLWQTLQALNF 116
Db 58 LKHPNIVRLDQVHNERKLYLVFEFLSQDLKKYMDSTPGSELPLHLIKSVLFLQLLQGVSF 117

Qy 117 CHIHNCIHRTKPNILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGD 175
Db 118 CHSHRVTHRDLPQNLLINELGAIKLADFGLARAFGVPLRTYTHEVTVLWYRAPEILLGS 177

Qy 176 TOYGSSVDIWAIGCVFAELLTGQPLWPGKSDVDQYLIIRTLGKLIPRHQSIFKSNQFFH 235
Db 178 KFYTTAVDIWSIGCIFAEMVTRKALFPGDSEIDQLFRIFRMLGT-----PSEDTWP 228

Qy 236 GIS-IPE-----PE-DMETLEEKFSVDVHPVALNFMKGCLKMNPDRLTCSQLLESSYFDS 288
Db 229 GVTQLPDYKGSFQKTRKGLESIVPNLEPEGRDLMLQLLYDPSQRITAKTALAHYPFSS 288

Qy 289 FQEAQIKRK 297
Db 289 PEPSPAARQ 297

RESULT 4
S40021
Protein kinase (EC 2.7.1.37) cdc2 homolog - slime mold (Dictyostelium discoideum)
C:Species: Dictyostelium discoideum
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 18-Jun-1999
C:Accession: S40021; S39071
R:Michaels, C.E.; Weeks, G.
submitted to the EMBL Data Library, August 1992
A:Description: The unicellular organism Dictyostelium discoideum possesses a highly re
A:Reference number: S40021
A:Accession: S40021
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-292 <MIC>
A:Cross-references: EMBL:L00652; NID:G167695; PIDN:AAA16056.1; PID:G167696
R:Michaels, C.; Weeks, G.
Biochim. Biophys. Acta 1179, 117-124, 1993
A:Title: The isolation from a unicellular organism, Dictyostelium discoideum, of a high
A:Reference number: S39071; MUID:94032415; PMID:8218353
A:Accession: S39071
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-18,'Y',20-188,'G',190-292 <MI2>
A:Cross-references: EMBL:L00652
C:Complex: In various organisms, cdc2 has been identified as a component of the M-phase
C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; cell cycle control; phosphoprotein; phosphotransferase; serine/threoni
F:2-254/Domain: protein kinase homology <KIN>
F:10-18/Region: protein kinase ATP-binding motif
F:33,51,126,128/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 34.6%; Score 574.5; DB 2; Length 292;
Best Local Similarity 37.5%; Pred. No. 1.e-20;
Matches 115; Conservative 64; Mismatches 93; Indels 35; Gaps 5;

Qy 1 MEKYEKLAKTGE GSYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQKH 60
Db 1 MEKYSKIEKLGE GTYGIWNKAKNRETGEIVALKRIRLDSEDEGVPCPTAIRISLLKELKH 60

Qy 61 PNLVNLIEVFRKRKMHLVFEYCDHTLLNELERNPNGVADGVKSVLWQTLQALNFC 120
Db 61 PNLVRLHDVITHTERKLTIVFEYLDQDLKKYLDCEGGEISKPTIKSFMVQLKGVAFCHDH 120

Qy 121 NCIHRTKPNILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGD 179
Db 121 RVLHRLDKPQNLLINRKGLKADFGLARAFGIPVRTYSHEVTVLWYRAPDVLMSRKYS 180

Qy 180 SSVDIWAIGCVFAELLTGQPLWPGKSDVDQYLIIRTLG-----KLIPRHQSIF- 228

Db 181 TPDIWSALCIFAEASGRPLFGSGTSDQLFRIFKILGTPNEESWPSITELPEYKTFDP 240
QY 229 -----KSGPFHGISIPEPEQMETLEKFSVDHPVALNFMKGLKONPDDRLTCSQLLES 293
Db 241 VHPAHQLSSIVHG-----LDEK-----GINLSSKMLQYDPNQRITAAALXKH 282
QY 284 SYFDSFQ 290
Db 283 PYFDGLE 289
RESULT 5
A40444
protein kinase (EC 2.7.1.37) cdc2 homolog A - maize
C;Species: Zea mays (maize)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A40444
R;Colasanti, J.; Tyers, M.; Sundaresan, V.
Proc. Natl. Acad. Sci. U.S.A. 88, 3377-3381, 1991
A;Title: Isolation and characterization of cDNA clones encoding a functional p34 (cdc2) h
A;Reference number: A40444; MUID:91195354; PMID:2014258
A;Accession: A40444
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-294 <COL>
A;Cross-references: GB:M60526
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: ATP; phosphotransferase
F;2-256/Domain: protein kinase homology <KIN>
F;10-18/Region: protein kinase ATP-binding motif
F;33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted
Query Match 34.4%; Score 571.5; DB 1; Length 294;
Best Local Similarity 37.9%; Pred. No. 1.5e-20;
Matches 118; Conservative 66; Mismatches 92; Indels 35; Gaps 9;
QY 1 MEKYEKLAKTGEISYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MEQYKVEKIGEGTYGVVYKALDKTANETIALKKIRLEQDEGEVFPSTAIRSILKEMNH 60
QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLN-----ELERNPNGVADGVKSVLWQTLOAL 114
Db 61 GNIVRLHDVVHSEKRIYLVFEYLDLKKFMDSCPEFAKNPT-----LIKSYLYQILRGV 115
QY 115 NFCHIHNCIHRDIPENILITKQ-GIIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELL 172
Db 116 AYCHSHRVLHRLDKPQNLLIDRRNALKLADFLARAFGIPVRTFTHEVTLWYRAPEIL 175
QY 173 VGDITQYGSVDIWAIGCVFAELLTGQPLWPGKSDVDQYLLIIRTLGKLIIPRHQSIFKSNG 232
Db 176 LGARQYSTPVDVMSVGCIFAEVMVNQKPLFPDGDSEIDELFKIFRILG--TPNEQS----- 227
QY 233 FFHGIS-IP-----EPEDMETLEKFSVDHPVALNFMKGLKONPDDRLTCSQLLE 282
Db 228 -WPGVSCLPDFKTAFFRWQAQDLATV---VPNLDPAGLDLLSKMLRYEPSKRITARQALE 283
QY 283 SSYFDSFQEAQ 293
Db 284 HEYFKDLEVVQ 294
RESULT 6
B40444
protein kinase (EC 2.7.1.37) cdc2 homolog B - maize (fragment)
C;Species: Zea mays (maize)
C;Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 28-Feb-1997
C;Accession: B40444
R;Colasanti, J.; Tyers, M.; Sundaresan, V.
Proc. Natl. Acad. Sci. U.S.A. 88, 3377-3381, 1991
A;Title: Isolation and characterization of cDNA clones encoding a functional p34 (cdc2) h
A;Reference number: A40444; MUID:91195354; PMID:2014258
A;Accession: B40444

A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-294 <COL>
A;Cross-references: GB:M60526
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: ATP; phosphotransferase
F;2-256/Domain: protein kinase homology <KIN>
F;10-18/Region: protein kinase ATP-binding motif
F;33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted
Query Match 34.4%; Score 571.5; DB 2; Length 294;
Best Local Similarity 37.6%; Pred. No. 1.5e-20;
Matches 117; Conservative 67; Mismatches 92; Indels 35; Gaps 9;
QY 1 MEKYEKLAKTGEISYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MEQYKVEKIGEGTYGVVYKALDKTANETIALKKIRLEQDEGEVFPSTAIRSILKEMNH 60
QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLN-----ELERNPNGVADGVKSVLWQTLOAL 114
Db 61 GNIVRLHDVVHSEKRIYLVFEYLDLKKFMDSCPEFAKNPT-----LIKSYLYQILRGV 115
QY 115 NFCHIHNCIHRDIPENILITKQ-GIIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELL 172
Db 116 AYCHSHRVLHRLDKPQNLLIDRRNALKLADFLARAFGIPVRTFTHEVTLWYRAPEIL 175
QY 173 VGDITQYGSVDIWAIGCVFAELLTGQPLWPGKSDVDQYLLIIRTLGKLIIPRHQSIFKSNG 232
Db 176 LGARQYSTPVDVMSVGCIFAEVMVNQKPLFPDGDSEIDELFKIFRVLG--TPNEQS----- 227
QY 233 FFHGIS-IP-----EPEDMETLEKFSVDHPVALNFMKGLKONPDDRLTCSQLLE 282
Db 228 -WPGVSCLPDFKTAFFRWQAQDLATI---VPNLDPAGLDLLSKMLRYEPSKRITARQALE 283
QY 283 SSYFDSFQEAQ 293
Db 284 HEYFKDLEVMVQ 294
RESULT 7
I5C474
protein kinase (EC 2.7.1.37) cdc2 [similarity] - goldfish
C;Species: Carassius auratus (goldfish)
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 05-May-2000
C;Accession: I5C474
R;Kajihira, H.; Yamashita, M.; Katsu, Y.; Nagahama, Y.
Dev. Growth Differ. 35, 647-654, 1993
A;Title: Isolation and characterization of goldfish cdc2, a catalytic component of mat
A;Reference number: I50474
A;Accession: I50474
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-302 <KAJ>
A;Cross-references: GB:D17758; NID:G471097; PIDN:BAAG4605.1; PID:G471098
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: ATP; phosphotransferase
F;2-256/Domain: protein kinase homology <KIN>
F;10-18/Region: protein kinase ATP-binding motif
F;33,51,128,130/Active site: Lys, Glu, Asp, Lys #status predicted
Query Match 34.3%; Score 570.5; DB 1; Length 302;
Best Local Similarity 38.6%; Pred. No. 1.7e-20;
Matches 120; Conservative 68; Mismatches 90; Indels 33; Gaps 7;
QY 1 MEKYEKLAKTGEISYGVVFKCRNKTSGQVAVKKF-VSEDDPVVKKIALREIRMLKQLK 59
Db 1 MDDYLKIEKIGEGTYGVVYKGRNKTGTQWAMKIRLESEEG-VPSTAVREISLKLQ 59
QY 60 HPNLVNLIEVFRKRKMHVFEYCDHTLLNELERNPNG--VADGVKSVLWQTLOALNFC 117
Db 60 HPNVRLDLVLMQESKLYLVFEFLSMDLKKYLDLSPSGQFMDPMLVKSILYQILEGILFC 119
QY 118 HIHNCIHRDIPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGD 176

```
Db      120 HCRRLVHRDLKPNQLLIDNKGVIKLDADFLARAFGVPVRVYTHEVWTLWYRAPEVLLGAS 179
QY      177 QYGSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIRTLG-----KLIPIRHQS 226
Db      180 RYSTPVDVWSIGTFAELATKKPLFHGDSEIDQLFRIFRKLTPNNEVWPDVESLPDYKN 239
QY      227 IF---KSNGFPHGISIPEPEDEMETLEBKPSDVHPVALNFMKGCCKNPNDRLTCSQLLES 283
Db      240 TFPKWKSG-----NLASTVKNLDKNGIDJLTKMLIYDPPKRIARQAYTH 284
QY      284 SYFDSFQEAQ 294
Db      285 PYFCDLDKST 295

RESULT 8
S22440
protein kinase (EC 2.7.1.37) cdc2 homolog 1 - rice
C:Species: Oryza sativa (rice)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 18-Jun-1999
C:Accession: S22440
R:Hashimoto, J.; Hirabayashi, T.; Hayano, Y.; Hata, S.; Ohashi, Y.; Suzuki, I.; Utsugi,
Mol. Gen. Genet. 233, 10-16, 1992
A:Title: Isolation and characterization of cDNA clones encoding cdc2 homologues from Ory
A:Reference number: S22440; MUID:92293101; PMID:1376401
A:Accession: S22440
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-294 <HAS>
A:Cross-references: EMBL:X60374; NID:G20342; PIDN:CAA42922.1; PID:G20343
C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein k
F:2-256/Domain: protein kinase homology <KIN>
F:10-18/Region: protein kinase ATP-binding motif
F:33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted
Query Match      34.2%; Score 568.5; DB 2; Length 294;
Best Local Similarity 37.6%; Pred. No. 2.1e-20;
Matches 117; Conservative 67; Mismatches 92; Indels 35; Gaps 9;

QY      1 MEKYEKLAKTGEISYGVVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIRMLKQLKH 60
Db      1 MEQYEKEEKIGESTYGVVYRARDKVTNETIALKKIRLEQEDGVPSTAIRSISLLKEMHH 60
QY      61 PNLVNLIEVFRKRKMHVFEYCDHTLNL-----ELERNPNGVADGVKSVLWQTLQAL 114
Db      61 GNIVRLHDVHSEKRIYLVFEYLDLKKFMDSCPEFAKNPT-----LIKSYLYQILRGV 115
QY      115 NFCHIHNCIHRDIKPENILITKO-GIHKIDFGFAQIL-IPGDAYIDYVATRWYRAPELL 172
Db      116 AYCHSHRVLRDLKPQNLIDRRRTNALKLADFLARAFGIPVATFTHEVVTWYRAPEIL 175
QY      173 VGDTCYGSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIRTLGKLIPIRHQSIFKSNG 232
Db      176 LGSRQYSTPVDVWSVGCIFAEMVWQKPLFPDSEIDCFKIFRVLG--TPNEQS----- 227
QY      233 FFHGI-SIP-----EPEDMETLEKFSDVHPVALNFMKGCCKNPNDRLTCSQLLE 282
Db      228 -WPGVSSLPDYKSAFPKWAQCLATI---VPTLDPAGLQLLSKMLRYEPNKRITARQALE 283
QY      283 SSYFDSFQEAQ 293
Db      284 HEYFKDLEMVQ 294

RESULT 9
A41227
protein kinase (EC 2.7.1.37) cdk2 - human
N:Alternate names: Egl homolog; protein kinase p34
C:Species: Homo sapiens (man)
C>Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 18-Jun-1999
C:Accession: A41227; S17873; S16520
```

```
R:Ninomiya-Tsujii, J.; Nomoto, S.; Yasuda, H.; Reed, S.I.; Matsumoto, K.
Proc. Natl. Acad. Sci. U.S.A. 88, 9006-9010, 1991
A:Title: Cloning of a human cDNA encoding a CDC2-related kinase by complementation of a
A:Reference number: A41227; MUID:92020980; PMID:1717994
A:Accession: A41227
A:Molecule type: mRNA
A:Residues: 1-298 <NIN>
A:Cross-references: GB:M68520; NID:G180177; PIDN:AAA35667.1; PID:G180178
R:Tsai, L.H.; Harlow, E.; Meyerson, M.
Nature 353, 174-177, 1991
A:Title: Isolation of the human cdk2 gene that encodes the cyclin A- and adenovirus E1F
A:Reference number: S17873; MUID:91367262; PMID:1653904
A:Accession: S17873
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-298 <TSA>
A:Cross-references: GB:X62071; NID:G312802; PIDN:CAA43985.1; PID:G312803
R:Edledge, S.J.; Spottiswood, M.R.
EMBO J. 10, 2653-2659, 1991
A:Title: A new human p34 protein kinase, CDK2, identified by complementation of a cdc2
A:Reference number: S16520; MUID:91330891; PMID:1714386
A:Accession: S16520
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-176, 'S', 178-298 <ELL>
A:Cross-references: EMBL:X61622; NID:G29848; PIDN:CAA43807.1; PID:G29849
C:Genetics:
A:Gene: GDB:CDK2
A:Cross-references: GDB:128984; OMIM:116953
A:Map position: 12q13-12q13
C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; cell cycle control; mitosis; phosphoprotein; phosphotransferase; serir
F:2-255/Domain: protein kinase homology <KIN>
F:10-18/Region: protein kinase ATP-binding motif
F:14,160/Binding site: phosphate (Thr) (covalent) #status predicted
F:15/Binding site: phosphate (Tyr) (covalent) #status predicted
F:33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted
Query Match      34.1%; Score 567; DB 2; Length 298;
Best Local Similarity 40.0%; Pred. No. 2.5e-20;
Matches 119; Conservative 60; Mismatches 99; Indels 19; Gaps 6;

QY      1 MEKYEKLAKTGEISYGVVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIRMLKQLKH 60
Db      1 MENFQVEKIGESTYGVVYKARNKLTGEVVALKIRLDTETEGVPSTAIRSISLLKELNH 60
QY      61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEARNP-NGVADGVKSVLWQTLQALNFCIH 119
Db      61 PNIKLLDVHTENKLYLVFEFLHQDLKKFMDASALTGIPLIKSYLFQLLQGLAFCHS 120
QY      120 HNCIHRDIKPENILITKQIKIKICDFGFAQIL-IPGDAYIDYVATRWYRAPELLVGDTCY 178
Db      121 HRVLRDLKPQNLINTEGAIKLADFLARAFGVVVRTYTHEVVTWYRAPEILLGCKYY 180
QY      179 GSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIRTLGKLIPIRHQSIFKSNGFFHGI- 237
Db      181 SFAVDIWSLGCIFAEMVTRRALFPDSEIDQLFRIFRTLGT-----PDEVVWPGVT 231
QY      238 SIPE--PEDMETLEKFSDVHPV---ALNFMKGCCKNPNDRLTCSQLLESYF 286
Db      232 SMPDYKPSFPKWARQDFSKVVPPLDEDEGRSLLSQMLHYDPNKRISAKAALAHPPF 286

RESULT 10
A37871
protein kinase (EC 2.7.1.37) cdk2 - African clawed frog
N:Alternate names: cell division control protein CDC2 homolog Egl
C:Species: Xenopus laevis (African clawed frog)
C>Date: 30-Apr-1991 #sequence_revision 30-Apr-1991 #text_change 18-Jun-1999
C:Accession: A37871; S15866; I51662; S14410
R:Paris, J.; Le Guellec, R.; Couturier, A.; Le Guellec, K.; Omilli, F.; Camonis, J.; Me
Proc. Natl. Acad. Sci. U.S.A. 88, 1039-1043, 1991
A:Title: Cloning by differential screening of a Xenopus cDNA coding for a protein high
```


Search completed: November 13, 2003, 14:21:17
Job time : 14.3233 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 14:12:56 ; Search time 9.51662 Seconds
(without alignments)
1556.584 Million cell updates/sec

Title: US-09-671-050-12

Perfect score: 1662

Sequence: 1 MEKYEKLAKTGEGSYGVVFX.....RKARNEGRRRRQQVPLXS 315

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1183.5	71.2	358	1	KKIA_HUMAN
2	720.5	43.4	1030	1	STK9_HUMAN
3	581	35.0	305	1	CDK3_HUMAN
4	574.5	34.6	292	1	CC2H_DICDI
5	571.5	34.4	297	1	CDK2_XENLA
6	570.5	34.3	302	1	CCC2_CARAU
7	568.5	34.2	294	1	CC21_ORYSA
8	567.5	34.1	303	1	CDC2_ORYCA
9	567	34.1	298	1	CDK2_HUMAN
10	566.5	34.1	294	1	CDC2_MAIZE
11	563	33.9	298	1	CDK2_CRIGR
12	561.5	33.8	294	1	CC2A_ARATH
13	561.5	33.8	303	1	CDC2_ORYLA
14	561	33.8	298	1	CDK2_RAT
15	559.5	33.7	303	1	CDC2_ORYCU
16	559.5	33.7	303	1	CDC2_ORYLU
17	559	33.6	298	1	CDK2_MESAU
18	559	33.6	302	1	CC21_XENLA
19	558.5	33.6	294	1	CDC2_CHERU
20	557	33.5	297	1	CDC2_RAT
21	556	33.5	297	1	CDC2_BOVIN
22	556	33.5	297	1	CDC2_HUMAN
23	556	33.5	302	1	CC22_XENLA
24	554	33.3	301	1	CC21_TRYBB
25	553	33.3	298	1	CDK2_CARAU
26	551.5	33.2	294	1	CDC2_VIGUN
27	551.5	33.2	311	1	CC23_TRYBB
28	551	33.2	288	1	CC2H_PLAFK
29	550.5	33.1	294	1	CDC2_VIGAC
30	548	33.0	297	1	CDC2_MOUSE
31	547.5	32.9	303	1	CDC2_CHICK
32	546.5	32.9	297	1	CDC2_DROME
33	545	32.8	292	1	CDK5_BOVIN
1	1183.5	71.2	358	1	Q00532 homo sapien
2	720.5	43.4	1030	1	Q76039 homo sapien
3	581	35.0	305	1	Q00526 homo sapien
4	574.5	34.6	292	1	P34117 dictyosteli
5	571.5	34.4	297	1	P23437 xenopus lae
6	570.5	34.3	302	1	P51958 carassius a
7	568.5	34.2	294	1	P29618 cryza sativ
8	567.5	34.1	303	1	Q9dga2 cryzias jav
9	567	34.1	298	1	P24941 homo sapien
10	566.5	34.1	294	1	P23111 zea mays im
11	563	33.9	298	1	O55076 cricetus
12	561.5	33.8	294	1	P24100 arabidopsis
13	561.5	33.8	303	1	Q9dgd3 cryzias lat
14	561	33.8	298	1	Q63699 rattus norv
15	559.5	33.7	303	1	Q9dga5 oryzias eur
16	559.5	33.7	303	1	Q9dgg8 oryzias luz
17	559	33.6	298	1	P48963 mesocricetu
18	559	33.6	302	1	P35567 xenopus lae
19	558.5	33.6	294	1	P93101 chenopodium
20	557	33.5	297	1	P39951 rattus norv
21	556	33.5	297	1	P48734 bos taurus
22	556	33.5	297	1	P06493 homo sapien
23	556	33.5	302	1	P24033 xenopus lae
24	554	33.3	301	1	P38973 trypanosoma
25	553	33.3	298	1	P43450 carassius a
26	551.5	33.2	294	1	P52389 vigna ungui
27	551.5	33.2	311	1	P54666 trypanosoma
28	551	33.2	288	1	Q07785 plasmodium
29	550.5	33.1	294	1	Q41639 vigna aconi
30	548	33.0	297	1	P11440 mus musculu
31	547.5	32.9	303	1	P13863 gallus gall
32	546.5	32.9	297	1	P23572 drosophila
33	545	32.8	292	1	Q02399 bos taurus

34	545	32.8	292	1	CDK5_RAT	Q03114 rattus norv
35	544	32.7	292	1	CDK5_HUMAN	Q0C535 homo sapien
36	543	32.7	292	1	CDK5_MOUSE	P49615 mus musculu
37	543	32.7	301	1	CC21_TRYCO	P54664 trypanosoma
38	541.5	32.6	294	1	CC22_MEDSA	Q05006 medicago sa
39	540	32.5	296	1	CDC2_DICDI	P34112 dictyosteli
40	536	32.3	292	1	CDK5_XENLA	P51166 xenopus lae
41	535.5	32.2	294	1	CC2A_ANTMA	Q38772 antirrhinum
42	533	32.1	302	1	CDC2_RANDY	Q9w739 rana dybows
43	533	32.1	305	1	PH85_YEAST	P17157 saccharomyc
44	530.5	31.9	346	1	CDK7_HUMAN	P50613 homo sapien
45	529	31.8	294	1	CDK5_DROME	P48609 drosophila

ALIGNMENTS

RESULT 1
KKIA_HUMAN
ID KKIA_HUMAN STANDARD; PRT; 358 AA.
AC Q00532;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Serine/threonine-protein kinase KKIALRE (EC 2.7.1.-) (Cyclin-dependent
DE kinase-like 1).
GN CDKL1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1];
RP SEQUENCE FROM N.A.
RX MEDLINE=92347325; PubMed=1639063;
RA Meyerson M., Enders G.H., Wu C.-L., Su L.-K., Gorka C., Nelson C.,
RA Harlow E., Tsai L.-H.;
RT "A family of human cdc2-related protein kinases.";
RL EMBO J. 11:2909-2917(1992).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; X66358; CAA47002.1; -.
EMBL; X66359; CAA47002.1; JOINED.
PIR; S23383; S23383.
HSSP; P24941; LHCL.
Genew; HGNC:1781; CDKL1.
GK; Q00532; -.
MIM; 603441; -.
GO; GO:0004693; F:cyclin-dependent protein kinase activity; TAS.
GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
GO; GO:0000074; P:regulation of cell cycle; TAS.
InterPro; IPR000719; Prot_kinase.
InterPro; IPR002290; Ser_thr_kinase.
Pfam; PF00069; pkinase; 1.
ProDom; PD000001; Prot_kinase; 1.
SMART; SM00220; S_TKc_1.
PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
Transferase; Serine/threonine-protein kinase; ATP-binding.
DOMAIN 5 288 PROTEIN_KINASE.
NP BIND 11 19 ATP (BY SIMILARITY).
BINDING 34 34 ATP (BY SIMILARITY).
ACT SITE 127 127 BY SIMILARITY.
SEQUENCE 358 AA; 41834 MW; 88344321F24B77C6 CRC64;


```
Query Match 71.2%; Score 1183.5; DB 1; Length 358;
Best Local Similarity 71.8%; Pred. No. 9.1e-82;
Matches 211; Conservative 41; Mismatches 41; Indels 1; Gaps 1;

QY 1 MEKYEKLAKTGEYSYGVVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIRMLKQLKH 60
DQ 2 MEKYEKIGKIGESYGVVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIRMLKQLKH 61
QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELRNPNGVADGVKSVLWQTLQALNFCIH 120
DQ 15-JUL-1999 (Rel. 38, Created)
DQ 15-JUL-1999 (Rel. 38, Last sequence update)
DQ 15-SEP-2003 (Rel. 42, Last annotation update)
DQ Serine/threonine protein kinase 9 (EC 2.7.1.37).
STK9.
OS Homo sapiens (Human).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98389628; PubMed=9721213;
RA Montini E., Andolfi G., Caruso A., Buchner G., Walpole S.M.,
RA Mariani M., Consalez G.G., Trump D., Ballabio A., Franco B.;
RT "Identification and characterization of a novel serine-threonine
RT kinase gene from the Xp22 region.";
RL Genomics 51:427-433(1998).
RN [2]
RP SEQUENCE OF 339-789 FROM N.A.
RA Krause S.W., Rehli M., Kreutz M., Schwarzfischer J., Paulauskis J.D.,
RA Andreesen J.D.;
RT "Differential screening leads to novel genetic markers of monocyte
RT to macrophage maturation.";
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphorylated protein.
CC -!- ENZYME REGULATION: COULD BE ACTIVATED BY PHOSPHORYLATION ON
CC TYROSINE AND THREONINE. PROBABLY BOTH PHOSPHORYLATIONS ARE
CC REQUIRED FOR ACTIVITY.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN BRAIN, LUNG, KIDNEY, PROSTATE,
CC OVARY, PLACENTA, PANCREAS AND TESTIS.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDK3 SUBFAMILY.
CC -!- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-1C IS THE INITIATOR.
CC -!- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC FRAMESHIFT IN POSITION 415.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC EMBL; Y15057; CAA75342.1; --
DR EMBL; X89059; CAA61445.1; ALT_FRAME.
DR HSSP; P24941; 1B38.
DR Genew; HGNC:11411; STK9.
DR GK; O76039; --
DR MIM; 300203; --
DR GO; GO:0004674; F:protein serine/threonine kinase activity; TAS.
DR GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW phosphorylation.
FT DOMAIN 13 297 PROTEIN KINASE.
FT NP_BIND 19 27 ATP (BY SIMILARITY).
FT BINDING 42 42 ATP (BY SIMILARITY).
FT ACT_SITE 135 135 BY SIMILARITY.
FT MOD_RES 169 169 PHOSPHORYLATION (ACTIVATES THE KINASE)
(POTENTIAL).
FT MOD_RES 171 171 PHOSPHORYLATION (ACTIVATES THE KINASE)
(POTENTIAL).
FT DOMAIN 784 789 POLY-LYS.
FT CONFLICT 339 340 HR -> GT (IN REF. 2).
FT CONFLICT 541 541 L -> W (IN REF. 2).
FT CONFLICT 731 764 MISSING (IN REF. 2).
SQ SEQUENCE 1030 AA; 115537 MW; 8A1C9C438610EF08 CRC64;

Query Match 43.4%; Score 720.5; DB 1; Length 1030;
Best Local Similarity 44.4%; Pred. No. 1.6e-46;
Matches 139; Conservative 69; Mismatches 98; Indels 7; Gaps 4;

QY 1 MEKYEKLAKTGEYSYGVVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIRMLKQLKH 60
DQ 10 MNKFEILGVGEGAYGVVFKCRHETHEIVAIKKFKDSEENEEVKETTLRELKMLRTLKQ 69
QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELRNPNGVADGVKSVLWQTLQALNFCIH 120
DQ 70 ENIVELKEAFRRRGKLYLVFEYVEKXNMELLEEMPNGVPPEKVKSYIQLIKAIHWCHKN 129
QY 121 NCIHRIKIPENILITKQGIKICDFGFAQLIPGD--AYTDYVATRWYRAPELLVGDYQY 178
DQ 130 DIVHRDIKPENLLISHNDVLKLCDFGFARNLSEGNANYTEYVATRWYRSPPELLLG-APY 188
QY 179 GSSVDIWAIGCVFAELLTGQPLWPGKSDVDQYLIIRTLGKLI PRHQSIFKSNFFHGHS 238
DQ 189 GKSVDMMWSVGCILGELSDGQPLFPGESEIDQLFTIQVGLPLPSEQMKLFYSNPRFHGLR 248
QY 239 IPEPEMETLEEKFSVDV-HPVALNFMKGCILKMNPDRLTCSQLLESSEYFDSFQEAQIKRK 297
DQ 249 FPAVNHPOSLERRYLGILNSVLLDMKNLKLDPADRYLTEQCLN---HPTFQQRLLDR 305
QY 298 ARNEGRNRRRQQV 310
DQ 306 SPSRSARPKPYHV 318

RESULT 3
CDK3 HUMAN
IC -CDK3 HUMAN STANDARD; PRT; 305 AA.
AC Q00526;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Cell division protein kinase 3 (EC 2.7.1.-).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```



```
Db 61 PNIVRLHDVHTERKLTLLVFEYLDQDLKKYLDECGGEISKPTIKSFYQLLKGVAFCHDH 120
Qy 121 NCIHRLDKPENILITKQGIKICDFGAQIL-IPGDYTDYVATRWYRAPELLVGDQY 179
Db 121 RVLHRDLKPONLLINRKGLKLAFLARAFGIPVPTYSHEVTLWYRAPEVLMSGSKYS 180
Qy 180 SSVDIWAIGCVFAELLTGQPLWPGKSDVDQYLIIRTLG-----KLIPRHQSIF- 228
Db 181 TPIDIWSALCIFAEMASGRPLPFGSGTSDQLFRIFKILGTPNEESWPSITELPEYKTDFF 240
Qy 229 -----KSNGFFHGISIPEPEMETLEEKFSOVHPVALNFMKGLKMNPDRLTCSQJLES 283
Db 241 VHPAQJSSIVHG-----JDEK-----GLNLLSKMLQYDPNGAITAAALKH 282
Qy 284 SYFDSFQ 293
Db 283 PYFDGLE 289

RESULT 5
CDK2 XENLA
ID -CDK2_XENLA STANDARD; PRT; 297 AA.
AC P23437;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cell division protein kinase 2 (EC 2.7.1.-) (cdk2 homolog EGI protein kinase).
GN EGI.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=91126051; PubMed=1704128;
RA Paris J., ie Guellec R., Couturier A., le Guellec K., Omilia F.,
RA Camonis J., Macneil S., Philippe M.;
RT "Cloning by differential screening of a Xenopus cDNA coding for a
RT protein highly homologous to cdc2.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:1039-1043(1991).
RN [2]
RP PHOSPHORYLATION OF THR-160.
RX MEDLINE=93345457; PubMed=8393783;
RA Poon R.Y.C., Yamashita K., Adamczewski J.P., Hunt T., Shuttleworth J.;
RT "The cdc2-related protein p40MO15 is the catalytic subunit of a
RT protein kinase that can activate p33cdc2 and p34cdc2.";
RL EMBO J. 12:3123-3132(1993).
CC -!- FUNCTION: PROBABLY INVOLVED IN THE CONTROL OF THE CELL CYCLE.
CC INTERACTS WITH CYCLINS A, B, OR E.
CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
CC THE ENZYME, WHILE PHOSPHORYLATION AT THR-160 ACTIVATES IT (BY
CC SIMILARITY).
CC -!- DEVELOPMENTAL STAGE: SYNTHESIZED IN UNFERTILIZED EGG, BUT NO
CC LONGER MADE IN THE EARLY EMBRYO.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; X14227; CAA32443.1; -
CC PIR; A37871; A37871.
CC HSSP; P24941; 1BUH.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR002290; Ser_thr_kinase.
```

```
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS0107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS0108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Cell cycle; Cell division; Mitosis; Phosphorylation.
FT DOMAIN 4 286 PROTEIN_KINASE.
FT NP_BIND 10 18 ATP (BY SIMILARITY).
FT BINDING 33 33 ATP (BY SIMILARITY).
FT ACT_SITE 127 127 BY SIMILARITY.
FT MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 160 160 PHOSPHORYLATION (BY CAK).
SQ SEQUENCE 297 AA; 33969 MW; EC30204FCB8D198C CRC64;

Query Match 34.4%; Score 571.5; DB 1; Length 297;
Best Local Similarity 38.1%; Pred. No. 5.8e-36;
Matches 17; Conservative 65; Mismatches 84; Indels 41; Gaps 6;

Qy 1 MEKYEKLAKTGESSYGVVFKCRNKTSGQVAVKKFVESEDSPVVKKIALREIRMLKQLKH 60
Db 2 MENFQKEKIGEGYGVVYKARNRETGEIVALKKIRLDTETEGVPSTAIRISLLKELNH 60
Qy 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELER-NPNGVADGVKSVLWOTLQALNFCHI 119
Db 61 PNIVKLLDVHTENKLYLVFEFLNQDLKKFMDRSNISGLALVKSYLFOLLQGLAFCHS 120
Qy 120 HNCIHRDIKPENILITKQGIKICDFGAQIL-IPGDYTDYVATRWYRAPELLVGDQY 178
Db 121 HRVLRDLKPONLLINSGAIIKLAFLARAFGVPVTRTFTEVTVLWYRAPELLGCKFY 180
Qy 179 GSSVDIWAIGCVFAELLTGQPLWPGKSDVDQYLIIRTLG-----KLIPRHQSIF 228
Db 181 STAVDIWSLGCIFAEMITRRALFPGDSEIDQLFRIFRLTGPDEVSWPGVTTMPDYKSTF 240
Qy 229 KSNGFFHGISIPEPEMETLEEKFSOVHPVALNFMKGLKMNPDRLTCSQJLESYFDS 288
Db 241 P-----KWIQDFSKVVP-----PLDEGGRDLLAQMLQ---YDS 271
Qy 289 PQEAQIK 295
Db 272 NKRIKSAK 278

RESULT 6
CDK2 CARC
ID -CDK2_CARAC STANDARD; PRT; 302 AA.
AC P51958;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cell division control protein 2 homolog (EC 2.7.1.-) (p34 protein kinase) (Cyclin-dependent kinase 1) (CDK1).
DE CDC2.
GN Carassius auratus (Goldfish).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Carassius.
OX NCBI_TaxID=7957;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Oocyte;
RA Kajiura H., Yamashita M., Katsu Y., Nagahama Y.;
RT "Isolation and characterization of goldfish cdc2, a catalytic
RT component of maturation-promoting factor.";
RL Dev. Growth Differ. 35:647-654(1993).
CC -!- FUNCTION: PLAYS A KEY ROLE IN THE CONTROL OF THE EUKARYOTIC CELL
CC CYCLE. IT IS REQUIRED IN HIGHER CELLS FOR ENTRY INTO S-PHASE AND
CC MITOSIS. COMPONENT OF THE KINASE COMPLEX THAT PHOSPHORYLATES THE
CC REPETITIVE CARBOXYL-TERMINUS OF RNA POLYMERASE II. CATALYTIC
CC COMPONENT OF MPF.
```

CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
CC THE ENZYME, WHILE PHOSPHORYLATION AT THR-16: ACTIVATES IT (BY
CC SIMILARITY).
CC -!- SUBUNIT: FORMS A STABLE BUT NON-COVALENT COMPLEX WITH CYCLIN B IN
CC MATURE OOCYTES.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D17758; BAA04605.1; -.
DR PIR: I50474; I50474.
DR HSSP: Q00534; 1B18.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00220; S_TKC; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Cell cycle; Cell division; Mitosis; Nuclear protein; Phosphorylation.
FT DOMAIN 4 287 PROTEIN_KINASE.
FT NP_BIND 10 18 ATP (BY SIMILARITY).
FT BINDING 33 33 ATP (BY SIMILARITY).
FT ACT_SITE 128 128 BY SIMILARITY.
FT MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 161 161 PHOSPHORYLATION (BY CAK) (BY SIMILARITY).
SQ SEQUENCE 302 AA; 34499 MW; 58DB812E19B311F5 CRC64;

Query Match 34.3%; Score 570.5; DB 1; Length 302;
Best Local Similarity 38.6%; Pred. No. 7e-36;
Matches 120; Conservative 68; Mismatches 90; Indels 33; Gaps 7;

QY 1 MEKYEKLAKTGEISYGVVFKCRNKTSGQVAVKKF-VESEDDPVVKKIALREIRMLKQLK 59
Db 1 MDDYLKIEKGEGTYGVVYKGRNKTGGVAVKKIRJEESEEG-VFSTAVREISLJKELQ 59

QY 60 HPNLVNLIEVFRKRKRKMHVFEYCDHTLLNELERNPG--VADGVKSVLWQTLQALNFC 117
Db 60 HPNVVRLLDVLNQESKLYLVFEFLSMCLKVKYLDSPSGQFNDPMCKVSKYLYQILEGILFC 119

QY 118 HIHNCIHRDIKPNILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELIVSDT 176
Db 120 HCRVVLHRDLKPQNLLIDNKGVIKLADEGLARAFGVPRVYTHEVVTLWYRAPEVLUGAS 179

QY 177 QYGSSVDIWAIGCVFAELLTGQPLWPKSKSDVQQLYLIIRTLG-----KLIPRHQS 226
Db 180 RYSTPVDVWSIGTIFAELATKKPLFHGDSIDQLFRIFRTLTGFENNEVWPVDESIPDYKN 239

QY 227 IF---KSNGFHGISIPEPEDMETLEEKFSDVHPFVALNFMKGLKMNFPDRLTCSQLLES 283
Db 240 TFPKWKSG-----NLASTVKNLDKNGIDLLTKYLIYDPPKRSARQAYTH 284

QY 284 SYFDSFQEAQI 294
Db 285 PYFDDLDKSTL 295

RESULT 7
CC21_ORYSA
ID CC21_ORYSA STANDARD; PRT; 294 AA.
AC P29618;
DT 01-APR-1993 (Rel. 25, Created)

DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Cell division control protein 2 homolog 1 (EC 2.7.1.-).
GN CDC2-1.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
[1]
SEQUENCE FROM N.A.
STRAIN=cv. Nipponbare;
XEDLINE=92293101; PubMed=1376401;
Hashimoto C., Hirabayashi T., Hayano Y., Hata S., Chashi Y.,
Suzuka I., Utsugi T., Toh-E A., Kikuchi Y.;
"Isolation and characterization of cDNA clones encoding cdc2
homologues from Oryza sativa: a functional homologue and cognate
variants.";
Mol. Gen. Genet. 233:10-16(1992).
CC -!- FUNCTION: PLAYS A KEY ROLE IN THE CONTROL OF THE EUKARYOTIC
CC CELL CYCLE. COMPONENT OF THE KINASE COMPLEX THAT PHOSPHORYLATES
CC THE REPETITIVE CARBOXYL-TERMINUS OF RNA POLYMERASE II.
CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X60374; CAA42922.1; -.
DR PIR: S22440; S22440.
DR HSSP: P24941; 1HCL.
DR Gramene; P29618; -.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00220; S_TKC; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Cell cycle; Cell division; Mitosis; Phosphorylation.
FT DOMAIN 4 287 PROTEIN_KINASE.
FT NP_BIND 10 18 ATP (BY SIMILARITY).
FT BINDING 33 33 ATP (BY SIMILARITY).
FT ACT_SITE 127 127 BY SIMILARITY.
FT MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 161 161 PHOSPHORYLATION (BY CAK) (BY SIMILARITY).
SQ SEQUENCE 294 AA; 34071 MW; 51322D93AEF4C131 CRC64;

Query Match 34.2%; Score 568.5; DB 1; Length 294;
Best Local Similarity 37.6%; Pred. No. 9.6e-36;
Matches 117; Conservative 67; Mismatches 92; Indels 35; Gaps 9;

QY 1 MEKYEKLAKTGEISYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MEQYEKEEKIGEGTYGVVYRARDKVTNETIALKKIRLEQDEGVPTAIREISLLKEMHH 60

QY 61 PNLVNLIEVFRKRKRKMHVFEYCDHTLLN-----ELEPNPGVADGVKSVLWQTLQAL 114
Db 61 GNVRLHDVHSEKRIYLVFEYLDLKKFKMDSCEFAKNPT-----LIKSYLQILRGV 115

QY 115 NFCHIHNCIHRDIKPNILITKQ-GIIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELL 172

Db 116 AYCHSHRVLRDLKPQNLLIDRRTNALKLADFLGLARAPGIPVRTFTHEVVTLWYRAPEIL 175
Qy 173 VGTQYGSVDIWAIGCVFAELLTGQPLWPGKSDVDQYLIIRTLGLKIPRHQSIFKSNG 232
Db 176 LGSROYSTPVDWMSVGCFAEMVNQKPLFPDSEIDELFKIFRVLG--TPNEQS----- 227
Qy 233 FFHGI-SIP-----EPDMETLEKFSVHPVAFNMFKGCLKMNPDRLTCSQJLE 282
Db 228 -WPGVSSLPDYKSAFPKWAQDLATI--VPTLDPAGLDLLSKMLRYEPNKRITARQALE 283
Qy 283 SSYFDSFQEAQ 293
Db 284 HEYFKDLENVQ 294

RESULT 8
CDC2_ORYJA STANDARD; PRT; 303 AA.
AC Q9DGA2; Q9DGA1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cell division control protein 2 homolog (EC 2.7.1.1) (p34 protein kinase)
DE kinase) (Cyclin-dependent kinase 1) (CDK1).
GN CDC2.
OS Oryzias javanicus (Japanese ricefish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha; Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
OX NCBI_TaxID=123683;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RA Yamashita M., Mita K.;
RT "cDNA cloning of Cdc2 and cyclin B in medaka species";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: PLAYS A KEY ROLE IN THE CONTROL OF THE EUKARYOTIC CELL CYCLE. IT IS REQUIRED IN HIGHER CELLS FOR ENTRY INTO S-PHASE AND MITOSIS. COMPONENT OF THE KINASE COMPLEX THAT PHOSPHORYLATES THE REPETITIVE CARBOXYL-TERMINUS OF RNA POLYMERASE II. CATALYTIC COMPONENT OF MPF (By similarity).
CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES THE ENZYME, WHILE PHOSPHORYLATION AT THR-161 ACTIVATES IT (BY SIMILARITY).
CC -!- SUBUNIT: FORMS A STABLE BUT NON-COVALENT COMPLEX WITH CYCLIN B IN MATURE OOCYTES (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AB050461; BAB17219.1; -.
EMBL; AB050462; BAB17220.1; -.
HSSP; Q00534; 1BI8.
InterPro; IPR000719; Prot_kinase.
InterPro; IPR002290; Ser_thr_kinase.
Pfam; PF00069; pkinase; 1.
ProDom; PD000001; Prot_kinase; 1.
SMART; SMC0220; S_TKc; 1.
PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
PROSITE; PS00108; PROTEIN_KINASE_S7; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Cell cycle; Cell division; Mitosis; Nuclear protein; Phosphorylation.
FT DOMAIN 4 287 PROTEIN KINASE.

FT NP BIND 10 18 ATP (BY SIMILARITY).
FT BINDING 33 33 ATP (BY SIMILARITY).
FT ACT SITE 128 128 BY SIMILARITY.
FT MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 161 161 PHOSPHORYLATION (BY CAK) (BY SIMILARITY).
FT VARIANT 49 49 V -> I.
FT VARIANT 245 245 M -> K.
FT VARIANT 269 269 I -> T.
FT VARIANT 276 276 I -> V.
FT VARIANT 299 299 C -> S.
SQ SEQUENCE 303 AA; 34691 MW; 439B8B012BE28D9C CRC64;

Query Match 34.1%; Score 567.5; DB 1; Length 303;
Best Local Similarity 37.9%; Pred. No. 1.2e-35;
Matches 117; Conservative 72; Mismatches 91; Indels 29; Gaps 7;

Qy 1 MEKYEKLAKTGEISYGVVFKCRNKTSGQVAVKKF-VESEDPVVVKIALREIRMLKQLK 59
Db 1 MEDYVKIEKIGETGVVYKGRHKSTGQVAVKKIRLESEEG-VPSTAVREVSLQELK 59
Qy 60 HENLVNLIIEVRRKRKXHLVFEYCDHTLLNELERNPNG--VADGVIKSVLWQTLQALNFC 117
Db 60 HENVVRLLDVLMQESRLYLIFEFLSMDLKKVLDIPSQGYMDPMLVKSPLYQILEGIYFC 119
Qy 118 H-HNCIHRDIKPENILITKQ--IKICDFGFAQIL-IPGDAYTQYVATRWYRAPELLVGDT 176
Db 120 HRRVLRDLKPQNLLIDNKGVIKLADFLGLARAFGVPVRVYTHEVVTLWYRAPEVLGSP 179
Qy 177 QYGSVDIWAIGCVFAELLTGQPLWPGKSDVDQYLIIRTLG-----KLIPRHQS 226
Db 180 RYSTPVDVWSTGTIFAELATKKPLFHGDSEIDQLFRIFRTLGTNPNDVWPDVESLPDYKN 239
Qy 227 IFKSNQFFHGISIPEDNE-TLEKFSVHPVAFNMFKGCLKMNPDRLTCSQJLESSY 285
Db 240 TF-----PKWMEGSLSSVMVKNLDKNCGLDLAKMLIYNPPKRISAREAMTHPY 286
Qy 286 FDSFQEAQI 294
Db 287 FDDLKSTL 295

RESULT 9
CDC2_HUMAN STANDARD; PRT; 298 AA.
AC P24941;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Cell division protein kinase 2 (EC 2.7.1.1) (p33 protein kinase).
GN CDK2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91330891; PubMed=1714386;
RA Elledge S.J., Spottswood M.R.;
RT "A new human p34 protein kinase, CDK2, identified by complementation of a cdc28 mutation in Saccharomyces cerevisiae, is a homolog of Xenopus Egl.";
RL EMBO J. 10:2653-2659(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91367262; PubMed=1653904;
RA Tsai L.-H., Harlow E., Meyerson M.;
RT "Isolation of the human cdk2 gene that encodes the cyclin A- and adenovirus E1A-associated p33 kinase.";
RL Nature 353:174-177(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92020980; PubMed=1717994;

RA Ninomiya-Tsuji J., Nomoto S., Yasuda H., Reed S.I., Matsumoto K.;
RT "Cloning of a human cDNA encoding a CDC2-related kinase by
RT complementation of a budding yeast cdc28 mutation.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:9006-9010(1991).
RN [4]
RP SEQUENCE FROM N.A.
RA Rieder M.J., Livingston R.J., Braun A.C., Montoya M.A., Chung M.-W.,
RA Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,
RA Schackwitz W.S., Sherwood J.K., Witrak L.A., Nickerson D.A.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBSJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [6]
RP PHOSPHORYLATION SITES.
RX MEDLINE=93010995; PubMed=1396589;
RA Gu Y., Rosenblatt J., O'Morgan D.O.;
RT "Cell cycle regulation of CDK2 activity by phosphorylation of Thr160
RT and Tyr15.";
RL EMBO J. 11:3995-4005(1992).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
RX MEDLINE=93288132; PubMed=8510751;
RA de Bondt H.L., Rosenblatt J., Cancarik J., Jones H.D.,
RA Morgan D.O., Kim S.-H.;
RT "Crystal structure of cyclin-dependent kinase 2.";
RL Nature 363:595-602(1993).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF COMPLEX WITH CYCLIN A.
RX MEDLINE=95356811; PubMed=7630397;
RA Jeffrey P.D., Russo A.A., Polyak K., Gibbs E., Hurwitz J.,
RA Massague J., Pavletich N.P.;
RT "Mechanism of CDK activation revealed by the structure of a
RT cyclinA-CDK2 complex.";
RL Nature 376:313-320(1995).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (2.33 ANGSTROMS) OF COMPLEX WITH L868276.
RX MEDLINE=96181476; PubMed=8610110;
RA de Azevedo W.F. Jr., Mulier-Dieckmann H.-J., Schulze-Gahmen U.,
RA Worland P.J., Sausville E., Kim S.-H.;
RT "Structural basis for specificity and potency of a flavonoid
RT inhibitor of human CDK2, a cell cycle kinase.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:2735-2740(1996).
RN [10]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF COMPLEX WITH CG2A AND KIP1.
RX MEDLINE=96300318; PubMed=8684460;
RA Russo A.A., Jeffrey P.D., Patten A.K., Massague J., Pavletich N.P.;
RT "Crystal structure of the p27Kip1 cyclin-dependent-kinase inhibitor
RT bound to the cyclin A-Cdk2 complex.";
RL Nature 382:325-331(1996).
RN [11]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF COMPLEX WITH CG2A.
RX MEDLINE=96313126; PubMed=8756328;
RA Russo A.A., Jeffrey P.D., Pavletich N.P.;
RT "Structural basis of cyclin-dependent kinase activation by
RT phosphorylation.";
RL Nat. Struct. Biol. 3:696-700(1996).
RN [12]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE=97075215; PubMed=8917641;
RA Schulze-Gahmen U., de Bondt H.L., Kim S.-H.;
RT "High-resolution crystal structures of human cyclin-dependent kinase
RT 2 with and without ATP: bound waters and natural ligand as guides for
RT inhibitor design.";
RL J. Med. Chem. 39:4540-4546(1996).
RN [13]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=97475219; PubMed=9334743;
RA Lawrie A.M., Noble M.E.M., Tunnah P., Brown N.R., Johnson L.N.,
RA Endicott J.A.;
RT "Protein kinase inhibition by staurosporine revealed in details of
RT the molecular interaction with CDK2.";
RL Nat. Struct. Biol. 4:796-801(1997).
RN [14]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF COMPLEX WITH CKS1.
RX MEDLINE=96182647; PubMed=8601310;
RA Bourne Y., Watson M.H., Hickey M.J., Holmes W., Rocque W., Reed S.I.,
RA Tainer J.A.;
RT "Crystal structure and mutational analysis of the human CDK2 kinase
RT complex with cell cycle-regulatory protein CksH1.";
RL Cell 84:863-874(1996).
RN [15]
RP X-RAY CRYSTALLOGRAPHY (2.05 ANGSTROMS).
RX MEDLINE=98342369; PubMed=9677190;
RA Gray N.S., Wodicka L., Thunnissen A.-M.W.H., Norman T.C., Kwon S.,
RA Espinoza F.H., Morgan D.O., Barnes G., Leclerc S., Meijer L.,
RA Kim S.H., Lockhart D.J., Schultz P.G.;
RT "Exploiting chemical libraries, structure, and genomics in the search
RT for kinase inhibitors.";
RL Science 281:533-538(1998).
CC -!- FUNCTION: PROBABLY INVOLVED IN THE CONTROL OF THE CELL CYCLE.
CC INTERACTS WITH CYCLIN A, D, OR E. ACTIVITY OF CDK2 IS MAXIMAL
CC DURING S PHASE AND G2.
CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
CC THE ENZYME, WHILE PHOSPHORYLATION AT THR-160 ACTIVATES IT.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X61622; CAA43807.1; -.
DR EMBL; X62071; CAA43985.1; -.
DR EMBL; M68520; AAA35667.1; -.
DR EMBL; AF512553; AAM34794.1; -.
DR EMBL; BC003065; AAH03065.1; -.
DR PIR; A41227; A41227.
DR PDB; 1FIN; 27-JAN-97.
DR PDB; 1HCK; 07-DEC-96.
DR PDB; 1HCL; 07-DEC-96.
DR PDB; 1AQL; 12-NOV-97.
DR PDB; 1JST; 11-JAN-97.
DR PDB; 1JSU; 29-JUL-97.
DR PDB; 1BUH; 09-SEP-98.
DR PDB; 1B38; 23-DEC-98.
DR PDB; 1B39; 23-DEC-98.
DR PDB; 1CKP; 13-JAN-99.
DR PDB; 1DI8; 28-JAN-03.
DR PDB; 1DM2; 31-MAY-00.
DR PDB; 1ELV; 10-MAY-01.

SQ SEQUENCE 303 AA; 34688 MW; 439B8D072BE48B9C CRC64;

Query Match 33.8%; Score 561.5; DB 1; Length 303;
Best Local Similarity 37.5%; Pred. No. 3.3e-35;
Matches 116; Conservative 72; Mismatches 92; Indels 29; Gaps 7;

Qy 1 MEKYEKLAKTGECSYGVVFKCRNKTSGQVAVKVF-VESEDDPVVKKIALREIRMLKQLK 59
Db 1 MEDYVKIEKIGEGTYGVVYKGRHKSTGQVAMKKIRLESEBEG-VPSTAVREVSLQLK 59

Qy 60 HPNLVNLIEVFRKRKMHVFEYCDHTLLNELERNPG--VADGVIKSVLWQTLQALNFC 117
Db 60 HPNVVRLLDVLMQESRLYLIFEFLSMDLKKYJDSIPSGQYMPMLVKSPLYQLLEGIVFC 119

Qy 118 HINCHIRDIKPENILITKQGIKICDFGFAQIL-IPSDAYTEYVATRWYRAPELJVGDT 176
Db 120 HRRVLHRDLKPONLLIDNKGVIKLADEFGLARAFGVPTVYTHEVVTLWYRAPEVLGSP 179

Qy 177 QYGSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLG-----KLIPRHQS 226
Db 180 RYSTPVDVWSTGTFAELATKKPLFHGDSEIDQLFRIFRTLTGTPNDVWPDVESLPDYK 239

Qy 227 IFKSNFFHGISIPEPEDVE-TLEKFSDVHPVALNFMKGCGLKMNPDRLTCSQLLESSY 285
Db 240 TF-----PKWKEGSLSMVKNLDKNGLDLLAKYLIYNPPKRISAREAMTHPY 286

Qy 286 FDSQFAQI 294
Db 287 FDDLKSTL 295

RESULT 14
CDK2_RAT
ID_CDK2_RAT STANDARD; PRT; 293 AA.
AC Q63699; O09136;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cell division protein kinase 2 (EC 2.7.1.-).
GN CDK2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Thyroid;
RX MEDLINE=95166553; PubMed=7862443;
RA Kotani S., Endo T., Kitagawa M., Higashi H., Onaya T.;
RT "A variant form of cyclin-dependent kinase 2 (Cdk2) in a malignant transformed rat thyroid (FRTL-Tc) cell line."
RL Oncogene 10:663-669(1995).
RN [2]
RP SEQUENCE OF 19-124 FROM N.A.
RX MEDLINE=96113578; PubMed=8673024;
RA Hosokawa Y., Yang M., Kaneko S., Taraka M., Nakashima K.;
RT "Synergistic gene expressions of cyclin E, cdk2, cdk5 and E2F-1 during the prolactin-induced G1/S transition in rat Nb2 pre-T lymphoma cells."
RL Biochem. Mol. Biol. Int. 37:393-399(1995).
CC -!- FUNCTION: PROBABLY INVOLVED IN THE CONTROL OF THE CELL CYCLE.
CC INTERACTS WITH CYCLINS A, D, OR E. ACTIVITY OF CDK2 IS MAXIMAL DURING S PHASE AND G2.
CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES THE ENZYME, WHILE PHOSPHORYLATION AT THR-160 ACTIVATES IT (BY SIMILARITY).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=CDK2-alpha;
CC IsoId=Q63699-1; Sequence=Displayed;
CC Name=CDK2-beta;
CC IsoId=Q63699-2; Sequence=Not described;
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

CC CDC2/CDKX SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
CC -----
CC EMBL; D28753; BAA05947.1; -;
CC EMBL; D63162; BAA09638.1; -;
CC HSSP; P24941; 1AQ1.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR002290; Ser_thr_kinase.
CC Pfam; PF00069; pkinase; 1.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00220; S_TKC; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding; Cell cycle;
KW Cell division; Mitosis; Phosphorylation; Alternative splicing.
FT DOMAIN 4 286 PROTEIN KINASE.
FT NP_BIND 10 18 ATP (BY SIMILARITY).
FT BINDING 33 33 ATP (BY SIMILARITY).
FT ACT_SITE 127 127 BY SIMILARITY.
FT MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 160 160 PHOSPHORYLATION (BY CAK) (BY SIMILARITY).
FT CONFLICT 79 79 V -> C (IN REF. 2).
FT CONFLICT 99 99 L -> I (IN REF. 2).
FT CONFLICT 124 124 L -> C (IN REF. 2).
SQ SEQUENCE 298 AA; 33887 MW; C8CB3ADCE9B97F88 CRC64;
Query Match 33.8%; Score 561; DB 1; Length 298;
Best Local Similarity 39.7%; Pred. No. 3.6e-35;
Matches 117; Conservative 61; Mismatches 99; Indels 18; Gaps 6;

Qy 1 MEKYEKLAKTGECSYGVVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIRMLKQLK 60
Db 1 MENFQVKEKIGEGTYGVVYKAKNKLGTGEWALKKIRLDTETEGVPSTAIRESLLKELNH 60

Qy 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELERNP-NGVADGVIKSVLWQTLQALNFC 119
Db 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELERNP-NGVADGVIKSVLWQTLQALNFC 119

Qy 120 HNCIHRDIKPENILITKQGIKICDFGFAQIL-IPSDAYTEYVATRWYRAPELJVGDTQY 178
Db 121 HRVLHRDLKPONLLINAEKSKADFGFLARAFGVPTVYTHEVVTLWYRAPEILGCKYY 180

Qy 179 GSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTGLKLIIPRHQSIFKSNFFHGI- 237
Db 181 STAVDWSLGCIFAEMVTRRALFPGDSEIDQLFRIFRTLTG-----PDEVVWPGVT 231

Qy 238 SIPE--PEDMETLEEKFSDVHPV---ALNFMKGCGLKMNPDRLTCSQLLESSYF 286
Db 232 SMPDYKPSFPKWARQDFSKVVPPLDEDEGRSLLSQMLHYDPNKRISAKAALAHPPF 286

RESULT 15
CDC2_ORYCU
ID_CDC2_ORYCU STANDARD; PRT; 303 AA.
AC Q9DGA5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cell division control protein 2 homolog (EC 2.7.1.-) (p34 protein kinase) (Cyclin-dependent kinase 1) (CDK1).
GN CDC2.
OS Oryzias latipes (Hymenoptera).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OV protein - protein search, using sw model

Run on: November 13, 2003, 14:16:06 ; Search time 29.9773 Seconds
(without alignments)
2711.600 Million cell updates/sec

Title: US-09-671-050-12
Perfect score: 1662
Sequence: 1 MEKYEKLAKTGECSYGVFK.....RKARNEGNNRRRQVLP LKS 315

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1206	72.6	352	11 Q8CEQ0	Q8ceq0 mus musculu
2	1027.5	61.8	353	5 Q9U2H1	Q9u2h1 caenorhabdi
3	1027.5	61.8	392	5 Q9VMN3	Q9vmn3 drosophila
4	967.5	58.2	566	6 Q9TTK0	Q9ttk0 oryctolagus
5	941.5	56.6	493	4 Q92772	Q92772 homo sapien
6	930.5	56.0	564	11 Q9QYI2	Q9qyi2 mus musculu
7	930.5	56.0	568	11 Q9QUK0	Q9quk0 mus musculu
8	928	55.8	329	11 Q9QYI1	Q9qyi1 mus musculu
9	784.5	47.2	455	4 Q9P114	Q9pil4 homo sapien
10	784.5	47.2	592	4 Q81VM4	Q81vw4 homo sapien
11	781	47.0	657	5 Q9BMG2	Q9bmg2 trypanosoma
12	770	46.3	353	11 Q8K134	Q8kl34 mus musculu
13	769.5	46.3	1106	5 Q9GRT9	Q9grt9 leishmania
14	767.5	46.2	457	11 Q8BKR2	Q8bkr2 mus musculu
15	764.5	46.0	585	11 Q8BL49	Q8bl49 mus musculu
16	763.5	45.9	457	11 Q9JM02	Q9jrm02 rattus norv

17	763.5	45.9	505	11 Q9JY01	Q9jrm01 rattus norv
18	759.5	45.7	595	11 Q8BLF2	Q8blf2 mus musculu
19	720.5	43.4	783	11 Q8BWI8	Q8bwi8 mus musculu
20	720.5	43.4	831	4 Q8IYC7	Q8iyc7 homo sapien
21	708	42.6	993	4 Q8WXQ5	Q8wxq5 homo sapien
22	702.5	42.3	578	11 Q8BVE0	Q8bve0 mus musculu
23	644	38.7	1104	13 Q9W6R6	Q9w6r6 fugu rubrip
24	641.5	38.6	997	4 Q9UJL6	Q9ujl6 homo sapien
25	579.5	34.9	294	10 Q9ZRI1	Q9zri1 triticum ae
26	564	33.9	288	5 Q9XZD6	Q9xzd6 plasmodium
27	562	33.8	288	5 Q96820	Q96820 plasmodium
28	560	33.7	288	5 Q96821	Q96821 plasmodium
29	555.5	33.4	300	3 Q13379	Q13379 pneumocysti
30	555.5	33.4	300	3 Q13380	Q13380 pneumocysti
31	555.5	33.4	300	5 Q15890	Q15890 toxoplasma
32	554.5	33.4	294	10 Q82666	Q82666 brassica na
33	554.5	33.4	294	10 Q40790	Q40790 pinus conto
34	552.5	33.2	300	5 Q44000	Q44000 toxoplasma
35	551	33.2	292	13 Q9DE44	Q9de44 brachydanio
36	550.5	33.1	294	10 Q43361	Q43361 picea abies
37	549.5	33.1	294	5 Q76541	Q76541 cryptospori
38	549.5	33.1	294	10 Q9FUR4	Q9fur4 nicotiana t
39	549.5	33.1	294	10 Q40484	Q40484 nicotiana t
40	549.5	33.1	297	10 Q8RX68	Q8rx68 arabidopsis
41	549	33.0	293	10 Q8GT22	Q8gt22 oryza sativ
42	549	33.0	298	5 Q27032	Q27032 theileria p
43	547.5	32.9	294	10 Q8L6T8	Q8l6t8 daucus caro
44	547.5	32.9	294	10 P93556	P93556 sesbania ro
45	547.5	32.9	300	5 Q17066	Q17066 asterina pe

ALIGNMENTS

RESULT 1

Q8CEQ0 ID Q8CEQ0 PRELIMINARY; PRT; 352 AA.
AC Q8CEQ0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Cyclin-dependent kinase-like 1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CS7BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573(2002).
DR EMBL; AK016781; BAC25497.1; -.
SQ SEQUENCE 352 AA; 41023 MW; 659F56C8080F35E0 CRC64;

Query Match	72.6%	Score 1206;	DB 11;	Length 352;
Best Local Similarity	68.9%	Pred. No. 2.2e-102;		
Matches 222;	Conservative 42;	Mismatches 48;	Indels 10;	Gaps 3;
QY	1	MEKYEKLAKTGECSYGVVFKCRNKTSGQWAVKKFVESEDDPVVKKIALREIRMLKQLKH	60	
DB	1	MEKYEKIGKIGESYGVVFKCRNRTDQIVAIKRFLETDQVVIKIALREIRMLKQLKH	60	
QY	61	PNLVNLIIEVFRKRKHLVFEYCDHTLLNELERNPNGVADGVKSVLWQTLQALNFCHIH	120	
DB	61	PNLVNLIIEVFRKRKHLVFEYCDHTLVHLELDRYQRGVPEPLVKNITWQTLQAVNFCCHKH	120	
QY	121	NCIHRDIXPENILITKQGIKICDPGFAQILI-PGDAYTDYVATRWYRAPPELLVGDTQYG	179	
DB	121	NCIHRDVKPENILITKQSAIKLDFGFARLLTGPGDYTDYVATRWYRSPPELLVGDTQYG	180	

QY	180	SSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIPRHQSIFKSNQFFHGISI	239
Db	181	PPVDVWAIGCVFAELLSGVPLWPGKSDVDQLYLIRKTLGDLIPRHQQVFVSMQYFSGVKI	240
QY	240	PEPEDMETLEKFSVHPVALNFMKGCLKQNPDDRLTCSQLLESSYFDSFQEA-----QI	294
Db	241	PDPEDMETLELKFPNISYSALGFLKGCJHMDPAERLTCEQLLQHPYFDSIREVGELTRQH	300
QY	295	KRKARNEGRRNRQ-----QVLP	312
Db	301	DKPARKTLRQSRKHLTGLQYLP	322
RESULT 2			
ID	Q9U2H1	PRELIMINARY;	PRT; 353 AA.
AC	Q9U2H1;		
DT	01-MAY-2000	(TREMBLrel. 13, Created;	
DT	01-MAY-2000	(TREMBLrel. 13, Last sequence update)	
DT	01-MAR-2003	(TREMBLrel. 23, Last annotation update)	
DE	Y42A5A.4	protein.	
GN	Y42A5A.4.		
OS	Caenorhabditis elegans.		
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;		
OC	Rhabditidae; Peloderinae; Caenorhabditis.		
OX	NCBI_TaxID=6239;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Steward C.A.;		
RL	Submitted (CCT-1998) to the EMBL/GenBank/DBSJ databases.		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=99069613; PubMed=9851916;		
RA	none;		
RT	"Genome sequence of the nematode C.elegans: A platform for		
RT	investigating biology.";		
RL	Science 282:2012-2018(1998).		
CC	-!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.		
DR	EMBL; AL032618; CAB63367.1; -.		
DR	HSSP; P24941; 1HCL.		
DR	WormPep; Y42A5A.4; CE20258.		
DR	InterPro; IPR000719; Prot_kinase.		
DR	InterPro; IPR002290; Ser_thr_pkinase.		
DR	Pfam; PF00069; pkinase; 1.		
DR	ProDom; PDC00001; Prot_kinase; 1.		
DR	SMART; SM00220; S_TKC; 1.		
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.		
DR	PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.		
DR	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.		
KW	ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.		
SQ	SEQUENCE 353 AA; 40665 MW; 7E32260D9849CAC9 CRC64;		
Query Match 61.8%; Score 1027.5; DB 5; Length 353;			
Best Local Similarity 58.8%; Pred. No. 5.4e-86;			
Matches 187; Conservative 59; Mismatches 63; Indels 9; Gaps 2;			
QY	1	MEKYEKLAKTGECSYGVVFKCRNKTSGQWVAVKKFVESEDQPVVKIALREIRMLKQLKH	60
Db	1	MDKYDRLSKLGESYGVVYCKNRDGTQVIAIKKFVETEDDPHIKKIALREIRMLKQLKH	60
QY	61	PNLVNLIIEVFRKRGKHLVFEYCDHTLLNELERNPNGVADGVIKSVLWQTLQALNFCHE	120
Db	61	QNLVGLIEVFKRNRKHLVFECDRTVLHELEKNPHGVNDELIIKKIYQLLEALKFCHSH	120
QY	121	NCIHRDIKPENILITKQGIKICDFGFAQLIPGDAYTCYVATRWYEAPELLVGTQVGS	180
Db	121	KCIHRDVKPENIFLTRNDQVKLGDFGFARIINTEMYTDYVATRWYRSPPELLVGDVQYGP	180
QY	181	SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIPRHQSIFKSNQFFHGISI	240
Db	181	PVDIWAIGCVYAEELLTGEALWPGRSIDIDLQLYHIRKTLGFLPRHISIFRTNQFFGLSIP	240

QY	241	EPEDMETLEKFSVHPVALNFMKGCLQKNPDDRLTCSQLLESSYFDSFQEAQIKRKARN	300
Db	241	EPEHLEPLPSKLPNASSAQLDFLQKCFEMSPDRRFSCSELMELHGIFSNW---ILRIQD	296
QY	301	EG-----RNRRRQQVLP	313
Db	297	ESTPTGLTSKRSPLYLP	314
RESULT 3			
Q9VMN3			
ID	Q9VMN3	PRELIMINARY;	PRT; 392 AA.
AC	Q9VMN3;		
DT	01-MAY-2000	(TREMBLrel. 13, Created)	
DT	01-MAY-2000	(TREMBLrel. 13, Last sequence update)	
DT	01-MAR-2003	(TREMBLrel. 23, Last annotation update)	
DE	CG7236	protein.	
GN	CG7236.		
OS	Drosophila melanogaster (Fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;		
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OC	Ephydroidea; Drosophilidae; Drosophila.		
OX	NCBI_TaxID=7227;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=BERKELEY;		
RX	MEDLINE=20196006; PubMed=10731132;		
RA	Adams M.D., Celniker S.E., Hoit R.A., Evans C.A., Gocayne J.D.,		
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,		
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,		
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,		
RA	Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,		
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,		
RA	Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,		
RA	Balliew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,		
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,		
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,		
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,		
RA	Cherry J.M., Cawley S., Dahike C., Davenport L.B., Davies P.,		
RA	de Pablos S., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,		
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,		
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,		
RA	Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,		
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guar P., Harris M.,		
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,		
RA	Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,		
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,		
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,		
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,		
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,		
RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,		
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,		
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacieb J.M.,		
RA	Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,		
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,		
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,		
RA	Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,		
RA	Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,		
RA	Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,		
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,		
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,		
RA	Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,		
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;		
RC	"The genome sequence of Drosophila melanogaster.";		
RL	Science 287:2185-2195(2000).		
CC	-!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.		
DR	EMBL; AE003611; AAF52279.1; -.		
DR	HSSP; P24941; 1HCL.		
DR	FlyBase; FBgn0031730; CG7236.		
DR	InterPro; IPR000719; Prot_kinase.		
DR	InterPro; IPR002290; Ser_thr_pkinase.		
DR	Pfam; PF00069; pkinase; 1.		
DR	ProDom; PDC00001; Prot_kinase; 1.		

```
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 392 AA; 44962 MW; C9627D077784E55D CRC64;

Query Match 61.8%; Score 1027.5; DB 5; Length 392;
Best Local Similarity 53.4%; Pred. No. 6.2e-86;
Matches 184; Conservative 65; Mismatches 61; Indels 5; Gaps 2;

QY 1 MEKYEKLAKTGEESYGWVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 PNLVNLIEVFRKRKQHLVFEYCDHTLLNELEARNPNPVGADGVKSVLWQTLQALNFCHIH 120
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MDRYEKLSRLGEGSYGVVYKCRDRETGALVAVKRFVESEDDPAIRKIALREIRLLKQLKH 60
61 PNLVNLIEVFRKRKQHLVFEYCDHTLLNELEARNPNPVGADGVKSVLWQTLQALNFCHIH 120
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 PNLVSLLEVFRKRKRLHLVFEFCELTVLHELHPQGCPEHSTKQICYQTLGVAYCHKQ 120
121 NCIHRRDIKPENILITKQGIKICDFGFAQILIPGDAYTDYVATRWYRAPELLVGDTOYG 179
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 GCLHRDIKPENILITKQGIKICDFGFAQILIPGDAYTDYVATRWYRAPELLVGDTOYG 180
181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIPRHQSIFKSNGFHGISIP 240
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
181 PVDVWAIGCLPAELVRGEALWPGRSVDQLYLIIRTLGCLLPRHIQIFGQNEYFKGITLP 240
QY 241 EPEDMETLEBKFDSDVHPVALNFMKGLKMPDDRLTCSQLLESSYFDSF--QEAQIK 295
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
241 VPPTLEPLEDKMPAKSQCNPLTIDFLKXCLDKDPTKRWSCCKLTKHSYFDDYIAKQRELE 300
QY 296 RKARNEGRNRRRQQV 310
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
301 HVNSLEAANLRQQQL 315

RESULT 4
Q9TTK0 PRELIMINARY; PRT; 566 AA.
AC Q9TTK0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Ser/Thr kinase KKIAMRE.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
CX NCBI_TaxID=9986;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Gomi H., Sun W., Finch C.E., Itohara S., Yoshimi K., Thompson R.F.;
RT "Learning induces a cdc2-related protein kinase, kkiamre."
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB029045; BAA88508.1; -.
DR HSSP; P24941; IHCL.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SMC0220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 566 AA; 64052 MW; 5D575955509C2EA9 CRC64;

Query Match 59.2%; Score 967.5; DB 6; Length 566;
Best Local Similarity 54.6%; Pred. No. 3.3e-80;
Matches 177; Conservative 67; Mismatches 65; Indels 15; Gaps 5;

QY 1 MEKYEKLAKTGEESYGWVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
```

```
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MEKYENLGLVGESYGMVMKCRNKDSGRIVAIAKKFLESDDDKMVKIAMREIKLLKQLRH 60
QY 61 PNLVNLIEVFRKRKQHLVFEYCDHTLLNELEARNPNPVGADGVKSVLWQTLQALNFCHIH 120
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 ENLVNLLEVCCKKRWYLVFEFVDHTLDDLELFPNGLDQVQVKYLFQIINGIGFCHSH 120
QY 121 NCIHRRDIKPENILITKQGIKICDFGFAQILIPGDAYTDYVATRWYRAPELLVGDTOYG 179
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
121 NIIHRDIKPENILVSQSGVVKLCDFGFARTLAAPGEVYTDYVATRWYRAPELLVGDVYK 180
QY 180 SSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIPRHQSIFKSNGFHGISI 239
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
181 KAVDVWAIGCLVTEMLMGEPLFPGSDIDQLYLIMRCLGNLIPRHQELFYKNPVFAGVRL 240
QY 240 PEPEDMETLEBKFDSDVHPVALNFMKGLKMPDDRLTCSQLLESSYF--DSF-----QEA 292
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
241 PEIKSEPLERRYPKLVSEVIDLAKKCLHVDPKRPFCAELLHHDFFQMDGFAERFSQEL 300
QY 293 QIK--RKARN-----EGRNRRRQQ 309
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
301 QMKVQKQDARNISLSKKSQNRKKEK 324

RESULT 5
Q92772 PRELIMINARY; PRT; 493 AA.
AC Q92772;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE P56 KKIAMRE protein kinase.
GN KKIAMRE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97152547; PubMed=9000130;
RA Taglienti C.A., Wysk M., Davis R.J.;
RT "Molecular cloning of the epidermal growth factor-stimulated protein
kinase p56 KKIAMRE."
RL Oncogene 13:2563-2574(1996).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; U35146; AAC50918.1; -.
DR HSSP; P24941; IHCL.
DR Genew; HGNC:1782; CDKL2.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 493 AA; 56018 MW; 5CC20A91CBF89EFE CRC64;

Query Match 56.6%; Score 941.5; DB 4; Length 493;
Best Local Similarity 53.1%; Pred. No. 6.6e-78;
Matches 172; Conservative 68; Mismatches 69; Indels 15; Gaps 5;

QY 1 MEKYEKLAKTGEESYGWVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MEKYENLGLVGESYGMVMKCRNKDTGRIVAIAKKFLESDDDKMVKIAMREIKLLKQLRH 60
QY 61 PNLVNLIEVFRKRKQHLVFEYCDHTLLNELEARNPNPVGADGVKSVLWQTLQALNFCHIH 120
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 ENLVNLLEVCCKKRWYLVFEFVDHTLDDLELFPNGLDYQVQVKYLFQIINGIGFCHSH 120
QY 121 NCIHRRDIKPENILITKQGIKICDFGFAQILIPGDAYTDYVATRWYRAPELLVGDTOYG 179
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

Db 121 NIIHRDIKPENILVSCGVVKLCJCFGARTLAAPGEVYTDVATRWYRAPELLVGDVKYG 180

QY 180 SSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLI PRHQSIKSNFFHGISI 239

Db 181 KAVDVWAIGCLVTEMFMGEPLFGDSIDQLYHIXVCLGNLIPRHQELFNKVPVAFAGVRL 240

QY 240 PEPEDMETLEEFSDVHPVALNFMKGCLKNPDDRLTCSQLLESSYF--DSF-----QEA 292

Db 241 PEIKEREPLERRYPKJSEVVIDLAKKCLHIDPKRPFCALLHHDFPQXDGFARFQSQEL 300

QY 293 QIK--RKARN-----EGRNRRRQQ 309

Db 301 QLKVQKDARNVSLSKKSQNRKKEK 324

RESULT 6

Q9QYI2

ID Q9QYI2 PRELIMINARY; PRT; 564 AA.

AC Q9QYI2

DT 01-MAY-2000 (TremBLrel. 13, Created)

DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)

DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)

DE Ser/Thr kinase KKIAMRE-gamma.

GN CDKL2 OR KKM.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN (1)

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6; TISSUE=Brain;

RA Sassa T., Gomi H., Sun W., Ikeda T., Thompson R.F., Itohara S.;

RT "The Murine KKIAMRE gene: variants, dual promoters, expression and chromosomal localization."

RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.

CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

DR EMBL; AB029066; BAA88428.1; -.

DR HSSP; P24941; IHCL.

DR MGD; MGI:1858227; Cdkl2.

DR InterPro; IPR000719; Prot_kinase.

DR InterPro; IPR002290; Ser_thr_pkinase.

DR Pfam; PF00069; pkinase; 1.

DR ProDom; PD000001; Prot_kinase; 1.

DR SMART; SM00220; S_TKC; 1.

DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.

DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.

KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.

SQ SEQUENCE 564 AA; 63640 MW; DACDEF630CCE82D5 CRC64;

Query Match 56.0%; Score 930.5; DB 11; Length 564;

Best Local Similarity 53.4%; Pred. No. 8.1e-77;

Matches 174; Conservative 66; Mismatches 69; Indels 17; Gaps 6;

QY 1 MEKYEKLAKTGEISYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60

Db 1 MEKYENLGLVGEISYGMVMKCRNKDSGRIVAIAKKFLESDDDKMVKKIAMREIKLLKQLRH 60

QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELERNPNGVADGVKSVLWQTLQALNFCIH 120

Db 61 ENLVNLLVCKKRWYLVFEFVDHTILDLLKLPNGLDYQVVQKYLFIINGIGFCHSH 120

QY 121 NCIHRDIPENILITKQGIKICDFGFAQIL-IPGDAYTDVATRWYRAPELLVGDVQY 179

Db 121 NIIHRDIKPENILVSCGVVKLCJCFGARTLAAPGEVYTDVATRWYRAPELLVGDVKYG 180

QY 180 SSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLI PRHQSIKSNFFHGISI 239

Db 181 KAVDIWAIGCLVIEMLMGQPLFGESDIDQLHHIMTCLGNLIPRHQELFYKNPVAFAGVRL 240

QY 240 PEPEDMET--LEEFSDVHPVALNFMKGCLKNPDDRLTCSQLLESSYF--DSF-----Q 290

Db 241 PEVKDAEAEPLERSYPKLPFAVISLAKKCLHIDPKRPFCADLLRHDFQMDGFAERFSQ 300

QY 291 EAQIK--RKARN-----EGRNRRRQQ 309

Db 301 ELQKIEKDARNNSLPKKSQNRKKEK 326

RESULT 7

Q9QUK0

ID Q9QUK0 PRELIMINARY; PRT; 568 AA.

AC Q9QUK0;

DT 01-MAY-2000 (TremBLrel. 13, Created)

DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)

DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)

DE Ser/Thr kinase KKIAMRE.

GN CDKL2 OR KKM.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN (1)

RP SEQUENCE FROM N.A.

RC STRAIN=129/SVJ; and C57BL/6;

RA Sassa T., Gomi H., Sun W., Ikeda T., Thompson R.F., Itohara S.;

RT "The Murine KKIAMRE gene: variants, dual promoters, expression and chromosomal localization."

RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.

CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

DR EMBL; AB029073; BAA88439.1; -.

DR EMBL; AB029068; BAA88439.1; JOINED.

DR EMBL; AB029069; BAA88439.1; JOINED.

DR EMBL; AB029070; BAA88439.1; JOINED.

DR EMBL; AB029071; BAA88439.1; JOINED.

DR EMBL; AB029072; BAA88439.1; JOINED.

DR EMBL; AB029065; BAA88427.1; -.

DR HSSP; P24941; IHCL.

DR MGD; MGI:1858227; Cdkl2.

DR InterPro; IPR000719; Prot_kinase.

DR InterPro; IPR002290; Ser_thr_pkinase.

DR Pfam; PF00069; pkinase; 1.

DR ProDom; PD000001; Prot_kinase; 1.

DR SMART; SM00220; S_TKC; 1.

DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.

KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.

SQ SEQUENCE 568 AA; 64055 MW; A43B75E2E9EB86C4 CRC64;

Query Match 56.0%; Score 930.5; DB 11; Length 568;

Best Local Similarity 53.4%; Pred. No. 8.2e-77;

Matches 174; Conservative 66; Mismatches 69; Indels 17; Gaps 6;

QY 1 MEKYEKLAKTGEISYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60

Db 1 MEKYENLGLVGEISYGMVMKCRNKDSGRIVAIAKKFLESDDDKMVKKIAMREIKLLKQLRH 60

QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELERNPNGVADGVKSVLWQTLQALNFCIH 120

Db 61 ENLVNLLVCKKRWYLVFEFVDHTILDLLKLPNGLDYQVVQKYLFIINGIGFCHSH 120

QY 121 NCIHRDIPENILITKQGIKICDFGFAQIL-IPGDAYTDVATRWYRAPELLVGDVQY 179

Db 121 NIIHRDIKPENILVSCGVVKLCJCFGARTLAAPGEVYTDVATRWYRAPELLVGDVKYG 180

QY 180 SSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLI PRHQSIKSNFFHGISI 239

Db 181 KAVDIWAIGCLVIEMLMGQPLFGESDIDQLHHIMTCLGNLIPRHQELFYKNPVAFAGVRL 240

QY 240 PEPEDMET--LEEFSDVHPVALNFMKGCLKNPDDRLTCSQLLESSYF--DSF-----Q 290

Db 241 PEVKDAEAEPLERSYPKLPFAVISLAKKCLHIDPKRPFCADLLRHDFQMDGFAERFSQ 300

QY 291 EAQIK--RKARN-----EGRNRRRQQ 309

RJ Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC041799; AAH41799.1; -
KW Kinase; Cyclin.
SQ SEQUENCE 592 AA; 67513 MW; 2B1AF08906E5B7697 CRC64;

Query Match 47.2%; Score 784.5; DB 4; Length 592;
Best Local Similarity 48.5%; Pred. No. 2.2e-63;
Matches 150; Conservative 61; Mismatches 91; Indels 7; Gaps 4;

QY 1 MEKYEKLAKTGEYSYGVVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIRMLKQLKH 60
DB 1 MEMYETLGKVGESYGTVMCKCHKNTGTQIVAIKIFYERPEQS-VNKIAMREIKFLKQFHH 59

QY 61 PNLVNLIEVFRKRKXHLVFEYCDHTLLNELEPNPVGADGVKSVLWQTLQALNFCIH 120
DB 60 ENLVNLIEVFRQKKXHLVFEFIDHTVLCEJQHYCHGLESKRLRKYLFQILRAIDYLHSN 119

QY 121 NCIHRIKPNILITKQGIKICDFGAQILIPGDAYTDYVATRWYRAPELLVGDTSYG 179
DB 120 NIHRDIKPNILVSQGITKLCDFGFARTLAPGDIYTDYVATRWYRAPELLVKTYSY 179

QY 180 SSVDIWAIGCVFAELLTGQPLWPGKSDVDQYLIIRTGLKLI PRHQSIKSNNGFFHGISI 239
DB 180 KPDVIWALGCMIEIEMATGNPYLPSSSDLLHLKIVLKVGNLSPHLQNI FSKSPIFAGVVL 239

QY 240 PEPEDMETLEKFSVDVHPVAFNMKGCLKMNPDRLTCSQLLESSYF--DSFQE---AQI 294
DB 240 PQVQHFKVARKKYPKLNGLLADIVHACLQIDPADRISSDLLHHEYFTRDGFIEKFNPEL 299

QY 295 KRKARNEGR 303
DB 300 KAKLQEA 308

RESULT 11
Q9BMG2 PRELIMINARY; PRT; 657 AA.
AC Q9BMG2;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-XAR-2003 (TREMBLrel. 23, Last annotation update)
DE Putative MAP kinase.
OS Trypanosoma brucei.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5691;
RN [1]
RP SEQUENCE FROM N.A.
RA Sarkar M., Matthews K.;
RT "Trypanosoma brucei: putative MAP kinase."
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AF326965; AAG49589.1; -
DR HSSP; P24941; 1B38.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00038; HLH_1; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 657 AA; 72116 MW; C93FA3C2D08F0619 CRC64;

Query Match 47.0%; Score 781; DB 5; Length 657;
Best Local Similarity 48.3%; Pred. No. 5.3e-63;
Matches 138; Conservative 67; Mismatches 81; Indels 0; Gaps 0;

QY 1 MEKYEKLAKTGEYSYGVVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIRMLKQLKH 60
DB 1 MDAYETLGMLEGTYGVVVKARHRTSRVAIKKYKQAECDJHVRKTSLEVRVLKQLRH 60

QY 61 PNLVNLIEVFRKRKXHLVFEYCDHTLLNELEPNPVGADGVKSVLWQTLQALNFCIH 120
DB 61 PNVIALLDVFRDGLYLVFEYVENTILQIEKKYGLSPDEVRTYTFQLLNGVSYCHAH 120

QY 121 NCIHRIKPNILITKQGIKICDFGAQILIPGDAYTDYVATRWYRAPELLVGDTSYG 180
DB 121 NIHRDVKPNILVSRDGVCLKCDFGARQLSCRGNYTEYVATRWYRAPELLVGDVSYGK 180

QY 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQYLIIRTGLKLI PRHQSIKSNNGFFHGISI 240
DB 181 AVDVWAIGCVFSELSDGQPLFPGSDLDQLSLIMRACGPVPOQMVSTFEHNALYRRVTFP 240

QY 241 EPEDMETLEKFSVDVHPVAFNMKGCLKMNPDRLTCSQLLESSYF 286
DB 241 SVDVEETLQQRFFTAASPWLEFLTSLRMDPVERPSC TALMSMAYF 286

RESULT 12
Q8K134
ID Q8K134 PRELIMINARY; PRT; 353 AA.
AC Q8K134;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Similar to serine/threonine kinase NKIATRE beta.
GN CDKL3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC028871; AAH28871.1; -
DR MGI; MGI:2388268; Cdk13.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00219; TyrKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Transferase.
SQ SEQUENCE 353 AA; 40619 MW; 7A645EE9FCAE2328 CRC64;

Query Match 46.3%; Score 770; DB 11; Length 353;
Best Local Similarity 45.7%; Pred. No. 2.4e-62;
Matches 149; Conservative 63; Mismatches 96; Indels 18; Gaps 4;

QY 1 MEKYEKLAKTGEYSYGVVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIRMLKQLKH 60
DB 1 MEMYETLGKVGESYGTVMCKCHKDTGRIVAIFKIFYE-KPEKSVNKIATREIKFLKQFRH 59

QY 61 PNLVNLIEVFRKRKXHLVFEYCDHTLLNELEPNPVGADGVKSVLWQTLQALNFCIH 120
DB 60 ENLVNLIEVFRQKKXHLVFEFIDHTVLDELOHYCHGLESKRLRKYLFQILRAIEYLHNN 119

QY 121 NCIHRIKPNILITKQGIKICDFGAQILIPGDAYTDYVATRWYRAPELLVGDTSYG 179
DB 121 NIHRDIKPNILVSQGITKLCDFGFARTLAPGDIYTDYVATRWYRAPELLVKTYSY 179

QY 180 SSVDIWAIGCVFAELLTGQPLWPGKSDVDQYLIIRTGLKLI PRHQSIKSNNGFFHGISI 239
DB 180 KPDVIWALGCMIEIEMATGHFPLPSSSDLLHLKIVLKVGNLTPHLHNIFSKSPIFAGVVL 239

QY 240 PEPEDMETLEKFSVDVHPVAFNMKGCLKMNPDRLTCSQLLESSYF--DSFQE---- 293
DB 240 PEPEDMETLEKFSVDVHPVAFNMKGCLKMNPDRLTCSQLLESSYF--DSFQE---- 293

Db 240 PQVQHPKTKARKKYPKLNGLLADIVHACLOIDPAERTSSTDLRHDYFTRDGFIEKKWTET 299

QY 294 -----LKRKARNEGRNRQQ 309

Db 300 KGRSSKSESJRPKGAKEMSQTRRSQ 325

RESULT 13

Q9GRT9 PRELIMINARY; PRT; 1106 AA.

AC Q9GRT9;

DT 01-MAR-2003 (TrEMBLrel. 16, Created)

DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Putative mitogen-activated protein kinase 6.

GN MPK6.

OS Leishmania mexicana.

OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.

OX NCBI_TaxID=5665;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MNYC/BZ/62/M379;

RA Wiese M., Wang Q., Goercke I.;

RT "Identification of mitogen-activated protein (MAP) kinases from

RT Leishmania mexicana.";

RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.

CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

DR EMBL; AJ293284; CAC07960.1; --

DR HSSP; P24941; 1B38.

DR InterPro; IPR000719; Prot_kinase.

DR InterPro; IPR002290; Ser_thr_pkinase.

DR Pfam; PF00069; pkinase; 1.

DR ProDom; PD000001; Prot_kinase; 1.

DR SMART; SM00220; S_TKC_1.

DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.

KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.

SQ SEQUENCE 1106 AA; 118884 MW; 24AC26F5C209923E CRC64;

Query Match 46.3%; Score 769.5; DB 5; Length 1106;

Best Local Similarity 48.3%; Pred. No. 1.2e-61;

Matches 145; Conservative 62; Mismatches 88; Indels 5; Gaps 2;

QY 1 MEKYEKLAKTGEYSYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60

Db 1 MEAYETLIGLGEPTYGVVVKARSRVTKLVAKRFKQTEQDEHVRKTSRSREVMQLQLQH 60

QY 61 PNLVNLIEVFRRRKRMHLVFEYCDHTLLNELERNPNGVADGVKSVLWQTLQALNFCCHI 120

Db 61 PNVIRLEDVFRREGKLYLVFEFIDHTLQLLESTTRGFHRHELRRYTYQLLRGIESCHNQ 120

QY 121 NCIHRIKPNENILITKQGIKICDFGFAQILIPGDAYTDYVATRWYRAPELLVGTQYGS 180

Db 121 NIHRDVKPNVLI DESGLLKLCDFGFAQTSAKGYTDYVATRWYRAPELLVGVAYGK 180

QY 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIIRTLGKLIPRHQSIFKSNGFHGISIP 240

Db 181 PVDVWALGMPFAELSDGQPLFGESDLDQLCLMQTCGPVPQRLVFIFMHNPLYNG-SFP 240

QY 241 EPEDMETLEEKFSDVHPVALNFMKGLKMNPDRLTCSQLLESSYF--DSFQ---EAQIK 295

Db 241 HTDILYTLKDRYHRESNDWIEFLSSCLHTDPAQRLTCTELMELPYFTRDGFRCRYAEELR 300

RESULT 14

Q8BKR2 PRELIMINARY; PRT; 457 AA.

AC Q8BKR2;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Serine/threonine kinase NKIATRE alpha homolog.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Body;

RX MEDLINE=22354683; PubMed=12466851;

RA The FANTOM Consortium,

RA the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs.";

RL Nature 420:563-573(2002).

DR EMBL; AK050990; BAC34488.1; --

SQ SEQUENCE 457 AA; 51900 MW; 6PCDD36F51152B61 CRC64;

Query Match 46.2%; Score 767.5; DB 11; Length 457;

Best Local Similarity 47.2%; Pred. No. 5.7e-62;

Matches 146; Conservative 64; Mismatches 92; Indels 7; Gaps 4;

QY 1 MEKYEKLAKTGEYSYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60

Db 1 MEMYETLGVGEGSYGTVMKCKKDKTGRIVAIFYE-KPEKSVNKIATREIKFLKQFRH 59

QY 61 PNLVNLIEVFRRRKRMHLVFEYCDHTLLNELERNPNGVADGVKSVLWQTLQALNFCCHI 120

Db 60 ENLVNLIEVFRQKKIHLVFEFIDHTVLDLQHYCHGLESKRRLKYLFOILRAIEYLHN 119

QY 121 NCIHRIKPNENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGTQYG 179

Db 120 NIIHRDIKPNENILVSQSGITKLCDFGFARTLAAPGDVYTDYVATRWYRAPELVLKDTSYG 179

QY 180 SSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIIRTLGKLIPRHQSIFKSNGFHGISI 239

Db 180 KPVDIWAIGCMIIEMATGHPFLRSSDLDLLHKIVLKVGNLTPHLHNIFSKSPIFAGVVL 239

QY 240 PEPEDMETLEEKFSDVHPVALNFMKGLKMNPDRLTCSQLLESSYF--DSFQ---AQI 294

Db 240 PQVQHPKTKARKKYPKLNGLLADIVHACLOIDPAERTSSTDLRHDYFTRDGFIEKFIPEL 299

QY 295 KRKARNEGR 303

Db 300 RAKLLQEAK 308

RESULT 15

Q8BL49 PRELIMINARY; PRT; 585 AA.

ID Q8BL49

AC Q8BL49;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Serine/threonine kinase NKIATRE alpha homolog.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Brain;

RX MEDLINE=22354683; PubMed=12466851;

RA The FANTOM Consortium,

RA the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs.";

RL Nature 420:563-573(2002).

DR EMBL; AK046394; BAC32701.1; --

SQ SEQUENCE 585 AA; 66530 MW; A1867A2A6CDF15A2 CRC64;

Query Match 46.0%; Score 764.5; DB 11; Length 585;

Best Local Similarity 46.9%; Pred. No. 1.5e-61;

Matches 145; Conservative 65; Mismatches 92; Indels 7; Gaps 4;

